

Disorder, entropy and harmonic functions

Itai Benjamini, Hugo Duminil-Copin, Gady Kozma and Ariel Yadin

November 22, 2011

Abstract

We study harmonic functions on random environments with particular emphasis on the case of the infinite cluster of supercritical percolation on \mathbb{Z}^d . We prove that the vector space of harmonic functions growing at most linearly is $(d+1)$ -dimensional almost surely. In particular, there are no non-constant sublinear harmonic functions (thus implying the uniqueness of the corrector). A main ingredient of the proof is a quantitative, annealed version of the Avez entropy argument. This also provides bounds on the derivative of the heat kernel, simplifying and generalizing existing results. The argument applies to many different environments, even reversibility is not necessary.

1 Introduction

Since the work of Yau [71] in 1975, where the Liouville property for positive harmonic functions on complete manifolds with non-negative Ricci curvature was proved, the structure of various spaces of harmonic functions have been at the heart of geometric analysis. Some years later, Yau conjectured that the space of polynomial growth harmonic functions of fixed order is always finite dimensional in open manifolds with non-negative Ricci curvature. Extensive literature has appeared on this conjecture and related problems. Understanding progressed quickly (Yau's conjecture was proved by Colding and Minicozzi [21]) and gave birth to many tools (see [56] for an introduction to the subject).

In algebraic setting bounded harmonic functions played a central role since the introduction of the Poisson boundary by Furstenberg [35, 36], see also the survey [70]. Recently, the geometric approach made a remarkable appearance in the algebraic realm when Kleiner proved that the space of harmonic functions with polynomial growth on the Cayley graph of a group with fixed polynomial volume growth is finite dimensional using the approach of [21]. He used this fact to provide a new proof of Gromov's theorem [51] (see [66] for a quantitative version of this theorem).

Another place where harmonic functions have played an important role recently is in the proof of the central limit theorem on random graphs. A central element in the proofs (see *e.g.* [67, 62, 14, 38]) is the construction of a harmonic function h on the cluster which is close to linear — the term $\chi(x) = h(x) - \langle x, v \rangle$ is called the corrector and once one shows that $\chi(x) = o(\|x\|)$, the proof may proceed.

The focus of this article is the case of random graphs. Classical tools of geometric analysis do not extend to this context in a straightforward way. Indeed, a random environment is not regular at the microscopic scale. In order to understand harmonic functions, one thus needs to control the macroscopic behavior of the environment. Let us take supercritical percolation as an example (see [39] for background and definitions).

For $p \in (0, 1)$, consider the random graph $G = (V(G), E(G))$ defined by $V(G) = V(\mathbb{Z}^d)$ and $E(G)$ being a random set containing each edge of \mathbb{Z}^d with probability p , independently of the other edges. It is classical that there exists $p_c(d) \in (0, 1)$ such that for $p < p_c(d)$, there is almost surely no infinite connected component (also called cluster), while for $p > p_c(d)$, there is a unique infinite cluster. When $p > p_c(d)$, we denote this cluster by ω .

Theorem 1. *Let $d \geq 2$ and let $p > p_c(d)$. Then with probability 1, the infinite cluster ω has no non-constant sub-linear harmonic functions.*

This immediately shows that the corrector χ is unique, as was conjectured by Berger and Biskup [14, Question 3].

In more regular settings, claims of this sort have been proved using the following strategy: try to show that two random walks starting at neighbors will couple before time n with probability bigger than $1 - Cn^{-1/2}$. This fact is classical in the case of the hypercubic lattice \mathbb{Z}^d where an explicit coupling can be exhibited. In the random context it is not clear how to construct an explicit coupling, but a number of approaches in the literature allow to construct a coupling indirectly. The known Gaussian heat kernel bounds (see (1) below) allow to construct a coupling that will fail with probability $n^{-\epsilon}$. Using also the central limit theorem already mentioned, one could improve this to $n^{-1/2+o(1)}$. Nevertheless, getting the precise $n^{-1/2}$ seems difficult with these approaches. The approach we will apply below not only gives the precise order $n^{-1/2}$, but the proof is also significantly simpler than those just suggested.

The proof uses an entropy argument similar to Avez [2] who showed that a Cayley graph satisfies the Liouville property if the entropy of random walk on it is sublinear. In fact the “if” here is an “if and only if” as was shown by Kaimanovich & Vershik [46] and, with a different approach, by Derriennic [30], but we will not need the other direction. Two extensions were known before for this result: it applies to random graphs [9] and it can be quantified [34, Section 5]. It turns out that both generalizations can be applied simultaneously. Further, Theorem 1 is but an example: the techniques work in great generality, even reversibility is not needed. Only stationarity of the walk and some weak (sub-)diffusivity are used. Precise assumptions are detailed below.

The environment as viewed from the particle. Before stating the results, we provide a precise definition of what we mean by ‘environment’. We are interested in environments which are somehow translational invariant. This notion extends the transitivity condition to the random context.

Consider a Markov chain P over some set V (formally a function $P : V \times V \rightarrow [0, 1]$ where $P(x, y)$ denotes the probability to move from x to y). We always assume our Markov chain is irreducible, *i.e.* that for any $v, w \in V$ there is an n such that $P^n(v, w) > 0$. A *rooted Markov chain* is a triplet (P, V, ρ) where $\rho \in V$ is some vertex that will be called

the *root vertex*. Two rooted Markov chains are considered isomorphic if there is a one-to-one map $\phi : V \rightarrow V'$ with $\phi(\rho) = \rho'$ and $P(x, y) = P'(\phi(x), \phi(y))$. An *environment as viewed from the particle*, abbreviated as simply *environment* is a random rooted Markov chain. Two environments are considered to have the same law if they can be coupled such that the resulting rooted Markov chains are isomorphic with probability 1.

Definition 2. An environment (P, V, ρ) is called *stationary* if it has the same law as (P, V, X_1) where X_1 is the position of the first step of the Markov chain P , starting from ρ , and independent of the random choice of (P, V, ρ) .

Stationary environments are very common and we provide ten examples in the end of Section 2. Most of these examples are embedded in \mathbb{Z}^d and for these we could have used a much simpler definition — there is no need to “couple the environments such that the resulting rooted Markov chains are isomorphic with probability 1”. One could just say that they have the same law considered as a random function $P : \mathbb{Z}^d \times \mathbb{Z}^d \rightarrow [0, 1]$, after shifting ρ to $\vec{0}$ (formally this is a stronger requirement, but all our examples embeddable into \mathbb{Z}^d have it). However, Examples 2.6, 2.8 and 2.10 are not embeddable into \mathbb{Z}^d , and we found no simpler definition that catches all examples of interest.

A very important subset of stationary environments is given by environments V with the structure of a weighted graph (with the weight being a positive function ν on every couple $(v, w) \in E$, and 0 on every couple $(v, w) \notin E$). P is given by

$$P(v, w) = \frac{\nu(v, w)}{\nu(v)} \quad \text{where } \nu(v) = \sum_x \nu(v, x).$$

These environments will be called *random stationary graphs*. This particular type of Markov chains is also commonly called *reversible*. The reversible case has a rich theory, see *e.g.* [1, 9] where one can also find many more examples. To clearly distinguish between the reversible and non-reversible case, random stationary graphs will be denoted by (G, ν, ρ) where G is the graph, ν is the weight function and ρ is the root.

The graph distance in G is denoted by $\mathbf{d}^G(\cdot, \cdot)$ and the ball of size r centered at x by $\mathbf{B}_x^G(r)$. We will also consider this distance in non-reversible setting, where it is simply the smallest n such that $P^n(x, y) > 0$. Since the distinction between annealed and quenched statements will be clear in the context, we will often drop the dependence on G in the notation. For instance, \mathbf{P}_x^G , $\mathbf{d}^G(\cdot, \cdot)$ and $\mathbf{B}_x^G(n)$ will become simply \mathbf{P}_x , $\mathbf{d}(\cdot, \cdot)$ and $\mathbf{B}_x(n)$. We collected all notations and conventions used in this paper in the last section of the introduction (page 8).

Non-constant harmonic functions with minimal growth. Let P be a Markov chain with state space V . Then a function $h : V \rightarrow \mathbb{R}$ is called harmonic if $h(X_n)$ is a martingale, or in other words, if

$$h(x) = \sum_y P(x, y)h(y) \quad \forall x.$$

As already mentioned, harmonic functions have had a number of important applications recently. Let us expand on the particular application in Kleiner’s proof of Gromov’s

theorem. It was known before Kleiner that to show Gromov's theorem it is enough to show that any group with polynomial volume growth has a non-trivial finite-dimensional representation. Kleiner showed that any group has a non-trivial linearly growing harmonic function, and that groups with polynomial growth have a finite dimension of those (even of polynomially growing functions). Since the group acts on harmonic functions on its Cayley graph by translations, this provides a finite dimension representation and proves Gromov's theorem. Shalom and Tao [66] showed that a quantitative version of Kleiner's proof can be performed. Further, they characterized the linearly growing harmonic functions — for groups with polynomial volume growth these are the non-constant harmonic functions with minimal growth [41, Theorem 6.1]). They showed (personal communication) that when the group is nilpotent, any such function must be a character of the group (or the sum of a character and a constant), in analogy to the Choquet-Deny Theorem [20, 60]. For virtually nilpotent groups this holds *mutatis mutandis*. We plan to analyze harmonic functions with minimal growth in the context of Cayley graphs, especially of wreath products, in a future paper.

We now return to the setting of this paper, i.e. of stationary random graphs. Using the entropy of the random walk, it is possible to bound from below the minimal growth of non-constant harmonic functions in terms of the rate of escape of the random walk. A particularly interesting case is the case of stationary environments with diffusive behavior, for which the bound is often sharp. A stationary environment (P, V, ρ) satisfies *diffusive or sub-diffusive behaviour (SDB)* if

$$(SDB) \quad \text{There exists } C > 0 \text{ such that } \mathbb{E}(\mathbf{d}(\rho, X_n)^2) \leq Cn \text{ for every } n.$$

where here and below \mathbb{E} is the average over both the environment and over the walk (the so-called annealed average). We may now state our main result:

Theorem 3. *Let (P, V, ρ) be a stationary environment such that $\mathbb{E}(|\mathbf{B}_\rho(n)|) \leq Cn^d$ for some constants $C, d < \infty$ independent of n . If (P, V, ρ) satisfies (SDB), then for almost every environment, there are no non-constant sublinear harmonic functions.*

As already stated, it applies to many different models, some of them significantly less well-understood than percolation. See a list of examples at the end of Section 2.

Whether (SDB) follows from polynomial growth in the reversible case is an interesting question. The Carne-Varopoulos bound [19, 69] gives that $\mathbf{E}_\rho(d(\rho, X_n)) \leq C\sqrt{n \log n}$ which would give (with the same proof as that of Theorem 3, see Theorem 3' in Section 2) that any stationary random graph with polynomial volume growth has no non-constant harmonic functions h with $h(x) \leq C\mathbf{d}(\rho, x)/\sqrt{\log \mathbf{d}(\rho, x)}$. Without stationarity the Carne-Varopoulos bound $\sqrt{n \log n}$ cannot be improved, as was shown by Barlow & Perkins [8]. Kesten gave a beautiful argument that a stationary random graph embedded in \mathbb{Z}^d satisfies (SDB), see *e.g.* [8, Section 2]. But it does not seem to apply just assuming polynomial growth.

The relation between entropy, harmonic functions and speed of the random walk holds for more general environments (for instance with larger growth). We defer to Section 2 for a more complete account of this question.

Polynomially growing functions. As in the case of manifolds, we are interested in the dimension of the space of harmonic functions with prescribed polynomial growth. Of course, one can encounter very different behavior depending on the environment (like in the deterministic case). Hence we will assume that our environments satisfy volume doubling and Poincaré inequality. In this section we do not assume stationarity, and hence we do not need a root for our graph. The object of interest is therefore simply a random weighted graph (G, ν) . Here is the precise formulation of our assumptions on the environment:

$(VD)_G$ (G, ν) satisfies the *volume doubling* property $(VD)_G$ if there exist $0 < \mathbf{C}_{VD} < \infty$ such that for every $\lambda < \infty$ and for every $x \in \mathbf{B}_\rho(\lambda n)$, $\nu(\mathbf{B}_x(2n)) \leq \mathbf{C}_{VD} \nu(\mathbf{B}_x(n))$ for n large enough, where $\nu(\mathbf{B})$ is the total weight of the edges in the ball \mathbf{B} .

$(P)_G$ (G, ν) satisfies the *Poincaré inequality* $(P)_G$ if there exists $\mathbf{C}_P < \infty$ such that for every $\lambda < \infty$, for every $x \in \mathbf{B}_\rho(\lambda n)$ and $f : \mathbf{B}_x(2n) \rightarrow \mathbb{R}$,

$$\sum_{y \in \mathbf{B}_x(n)} (f(y) - \bar{f}_{\mathbf{B}_x(n)})^2 \nu(y) \leq \mathbf{C}_P n^2 \sum_{(y,z) \in E(\mathbf{B}_x(2n))} |f(y) - f(z)|^2 \nu(y, z)$$

for n large enough, where

$$\bar{f}_{\mathbf{B}_x(n)} = \frac{1}{\nu(\mathbf{B}_x(n))} \sum_{y \in \mathbf{B}_x(n)} f(y) \nu(y).$$

\mathbf{C}_{VD} and \mathbf{C}_P may depend on the random choice of (G, ν) (though we do not have any interesting example which actually uses this freedom). The minimal n from which the properties hold may depend both on the environment and on λ .

These properties are classical in geometric analysis. They go back to the theory developed by De Giorgi, Nash and Moser [63, 64, 65, 26] in the fifties and sixties for uniformly elliptic second-order operators in divergence form. In the deterministic context, they imply the Harnack principle and Gaussian bounds for the heat kernel. The versions above are tailored for the random case: they take into consideration that in most examples of interest these properties do not hold from every point since some unusual points always exist. So we require that they hold for balls which are not too far from our root ρ , relative to their size. This is reminiscent of Barlow's good and very good balls [4], but our requirements are much weaker, we only need the properties to hold for “macroscopic balls”, balls whose distance to ρ is proportional to their size.

Let us remark on the appearance of the number 2 in $\mathbf{B}_x(2n)$ in both properties. For the volume doubling property it is clear that these properties are equivalent for all choices bigger than 1 *i.e.* if one were to define a “3-volume doubling property” then it would be equivalent to the “2-volume doubling property” defined above, though perhaps with different \mathbf{C}_{VD} and minimal n . The same holds for the Poincaré inequality, under the assumption of volume doubling. This is well known in deterministic settings, see *e.g.* [45, §5], and the proof carries over to the random case with no change.

With these definitions we can state the following easy but, we believe, conceptually important theorem:

Theorem 4. *Let (G, ν) be a random weighted graph (not necessarily stationary). If (G, ν) satisfies $(VD)_G$ and $(P)_G$, then for every $k > 0$, the space of harmonic functions with $h(x) \leq \mathbf{d}(\rho, x)^k$ for all x large enough, is finite dimensional.*

Further, the bound on the dimension depends only on \mathbf{C}_{VD} and \mathbf{C}_P .

This theorem represents a randomized discrete version of Yau conjecture (except that the Poincaré inequality must be assumed since it is not automatically satisfied). The proof of this theorem follows the existing path of [21, 27, 51, 66, 68]. Let us stress again that the interesting part is that it requires only *macroscopic* volume growth and Poincaré inequality: the definitions of $(VD)_G$ and $(P)_G$ only examine balls of radius n inside $\mathbf{B}_\rho(\lambda n)$ for some finite λ .

Since the dimension depends only on \mathbf{C}_{VD} and \mathbf{C}_P , then in particular, if these constants are not random, neither is the bound. Thus, for example, in supercritical percolation there is a constant A (depending only on the dimension d and the probability p) such that $\mathbf{C}_{VD} \leq A$ and $\mathbf{C}_P \leq A$ almost surely (the minimal n is the only quantity which really changes between configurations). Hence for each k there is a number D_k such that the dimension of harmonic functions of growth smaller than $|x|^k$ is smaller than D_k , almost surely.

Linearly growing functions. In the special case of environments which are modifications of \mathbb{Z}^d , we can compare the dimension of harmonic functions with a prescribed growth to the dimension of harmonic functions on \mathbb{Z}^d . The simplest perturbation of \mathbb{Z}^d is the supercritical cluster of percolation. We prove the following theorem:

Theorem 5. *For $p > p_c(d)$, let ω be the unique infinite component of percolation on \mathbb{Z}^d . Then, the dimension of the vector space of harmonic functions with growth at most linear on ω is equal to $d + 1$ almost surely.*

This theorem must be understood as a first step towards a bigger goal, which would be to compute the dimension of all spaces with prescribed growth.

The properties of the supercritical percolation cluster that we harness in this proof are quite general. First we use the d -dimensional volume growth and the Poincaré inequality $(P)_\omega$, proved (in stronger form) by Barlow [3], as well as the Gaussian bounds which Barlow concludes from these. But the main ingredient of the proof is an invariance principle [67, 14, 62]. All these properties witness the close relation between macroscopic properties of the supercritical percolation cluster and \mathbb{R}^d . In some sense, it confirms the fact that this cluster is an approximation of \mathbb{Z}^d .

Heat kernel estimates. Classically it is known that the kernels of symmetric diffusions have some Hölder regularity. By ‘classically’ we refer to [63, 64, 65, 26]. In random environments, few results are known on Hölder behavior: Conlon and Naddaf [23] and Delmotte and Deuschel [29] treated the case of random conductance with a uniform ellipticity condition, see also [38]. The entropy techniques developed for the proof of theorem 1 allow to give a very short proof that the space derivative exists. Moreover, it applies in a very general context. We present the case of percolation.

Theorem 6. Let \mathbb{P}_p be the measure of the infinite cluster of percolation (denoted ω) on \mathbb{Z}^d . There exist $C_3, C_4 > 0$ such that for every $n > 0$ and x, x', y at distance less than n of 0, $x \sim x'$,

$$\mathbb{E}_p \left[\left(\mathbf{p}_n(x, y) - \mathbf{p}_{n-1}(x', y) \right)^2 \mathbf{1}_{\{y, x \sim x' \in \omega\}} \right] \leq \frac{C_3}{n^{d+1}} \exp[-C_4|x - y|^2/n],$$

where $\mathbf{p}_n(y, x) := \mathbf{P}_y(X_n = x)$ and X_n is the random walk on ω .

Estimates for the heat kernel itself (*i.e.* not for the derivative) are well understood, and are known as Gaussian estimates (*GE*). Heuristically, Gaussian estimates are bounds of the form

$$\frac{C_1}{n^{d/2}} \exp[-C_2|x - y|^2/n] \leq \mathbf{P}_x[X_n = y] \leq \frac{C_3}{n^{d/2}} \exp[-C_4|x - y|^2/n].$$

A few caveats are in place, though. The lower bound cannot hold if there is any kind of periodicity (as in \mathbb{Z}^d or in subgraphs of it, such as supercritical percolation). One should talk about continuous time random walk, lazy random walk, or replace $\mathbf{P}_x[X_n = y]$ with $\mathbf{P}_x[X_n = y] + \mathbf{P}_x[X_{n+1} = y]$. Further, the lower bound does not hold for x and y extremely far away — if $|x - y| > t$ then the probability is just zero (in the simple random walk case).

In the case of the infinite cluster of supercritical percolation, these bounds were obtained for continuous time random walk in [4]. They also hold for simple random walk, most of the details are filled in in [7]. Again, one should be careful, as (with small probability) the environment in the neighborhood of ρ might be atypical, breaking these estimates for small t . Hence the formulation is as follows. There exist strictly positive constants C_1, C_2, C_3 and C_4 such that for almost every environment ω there exist random variables $n_x(\omega), x \in \mathbb{Z}^d$ so that for every $x, y \in \omega$ and $n > n_x(\omega), |x - y|$.

$$\frac{C_1}{n^{d/2}} \exp[-C_2|x - y|^2/n] \leq \mathbf{P}_x[X_n = y] + \mathbf{P}_x[X_{n+1} = y] \leq \frac{C_3}{n^{d/2}} \exp[-C_4|x - y|^2/n]. \quad (1)$$

Moreover, the random variables $n_x(\omega)$ satisfy a stretched exponential estimate *i.e.*

$$\mathbb{P}_p(x \in \omega, n_x(\omega) \geq s) \leq ce^{-cs^\varepsilon} \quad (2)$$

for some $\varepsilon > 0$.

For the proof of Theorem 6 we only need the upper bound in (1). For the proof of Theorem 5 we will also need the lower bound, but only in the regime $|x - y| \approx \sqrt{n}$ *i.e.* in the regime where the probabilities are $\approx n^{-d/2}$.

Organization of the paper. In the next section, we study the notion of mean entropy of random walks on a stationary random graph to bound the total variation between random walks starting at neighbors. We deduce Theorem 3. Section 3 contains the proof that $(VD)_G$ and $(P)_G$ imply that the space of harmonic functions of prescribed polynomial growth is finite dimensional, *i.e.* Theorem 4. Section 4 deals with the example of the supercritical percolation cluster and analyzes the space of linearly growing harmonic functions. It is completely independent of Section 3. Section 5 contains the proof of Theorem 6. Section 6 regroups all the open questions.

Notations. To make the distinction between the reversible and non-reversible case clear, we call the general case “Markov chain” and denote it by (P, V) , where V is the space and $P : V \times V \rightarrow [0, 1]$ are the transition probabilities, $P(x, y)$ being the probability to move from x to y . We often write P^n which we interpret as a matrix power — of course, $P^n(x, y)$ is also the probability that a random walk starting from x will be at y after n steps.

The reversible case is called “graph” and denoted by (G, ν) where G is a graph and ν are weights on the edges *i.e.* $\nu : E(G) \rightarrow [0, \infty)$. Here and below we denote by $E(G)$ the set of edges of the graph G , and for a set of vertices S we denote by $E(S)$ the set of edges between the vertices of S . The notation $x \sim y$ for two vertices will mean that they are neighbors in the graph.

We also consider ν as a measure. So for a vertex x we will denote $\nu(x) = \sum_{y \sim x} \nu(x, y)$ while for a set of vertices S we will denote $\nu(S) = \sum_{x \in S} \nu(x)$. Note that edges between two vertices of S are counted twice in this sum.

For a fixed graph or Markov chain we denote by \mathbf{E} the expectation with respect to the random walk on that fixed graph. When we want to note the starting point of the random walk we will use subscripts, *i.e.* \mathbf{E}_ρ . \mathbb{E} is used to denote the expectation with respect to both the environment and the random walk (the “annealed” average). Similarly, bold letters will usually denote “quenched” objects, *i.e.* objects related to an instance G of the environment. The quantity $\mathbf{d}(x, y)$ will denote the graphical distance between two vertices x and y of G , *i.e.* the length of the shortest path in G between x and y , or, in the non-reversible setting, the minimal n such that $P^n(x, y) > 0$. $\mathbf{B}_x(r)$ will denote the ball $\{y : \mathbf{d}(x, y) \leq r\}$. \mathbf{c}_i will denote constants which depend on the environments G , while constants of the form C_i are uniform in the environment. We will occasionally write just \mathbf{c} or C for a constant — different appearances of \mathbf{c} or C might be different constants.

Finally, for a random variable X we denote by $\mathcal{L}(X)$ the law of X , *i.e.* the measure on the space of values of X induced by it. If \mathcal{E} is some event then we will denote by $\mathcal{L}(X|\mathcal{E})$ the law of X conditioned on \mathcal{E} happening. For a set E we will denote by $|E|$ the cardinality of E .

2 The Entropy Argument

The connection between entropy and random walks was first exhibited by Avez [2] and then made famous in a celebrated paper of Kaimanovich and Vershik [46] (see also Derriennic [30]). For any discrete variable X the entropy is defined by

$$H(X) = \sum_x \phi(P(X = x)) \quad \phi(t) = -t \log t, \phi(0) = 0.$$

Consider a stationary environment (P, V, ρ) with law \mathbb{P} . Conditionally on (P, V, ρ) , define the *entropy* of the random walk started at x at times n, m by

$$\mathbf{H}_n^m(P, V, \rho) = H(X_n, X_m) = \sum_{x, y \in V} \phi(\mathbf{P}_\rho(X_n = x, X_m = y))$$

When $n = m$, we simply denote $\mathbf{H}_n^n(P, V, \rho)$ by $\mathbf{H}_n(P, V, \rho)$. In the random context, we define the *mean entropy* (see [9]) by

$$H_n^m = \mathbb{E}[\mathbf{H}_n^m(P, V, \rho)] \quad \text{and} \quad H_n = \mathbb{E}[\mathbf{H}_n(P, V, \rho)].$$

There are many ways of measuring the distance between two probability measures μ and ν on some set V , the most standard being the total variation $\|\mu - \nu\|_{TV} := \sum_{x \in V} |\mu(x) - \nu(x)|$. In this article, we will use a less standard one. Define $\Delta(\mu, \nu)$ by the formula

$$\Delta(\mu, \nu) := \left[\sum_{x \in G} \frac{(\mu(x) - \nu(x))^2}{\mu(x) + \nu(x)} \right]^{1/2}. \quad (3)$$

Estimating the distance using Δ is stronger than via the total variation: by Cauchy-Schwartz,

$$\begin{aligned} \|\mu - \nu\|_{TV} &= \sum_x |\mu(x) - \nu(x)| = \sum_x \sqrt{\mu(x) + \nu(x)} \frac{|\mu(x) - \nu(x)|}{\sqrt{\mu(x) + \nu(x)}} \leq \\ &\leq \sqrt{\left(\sum_x \mu(x) + \nu(x) \right) \left(\sum_x \frac{(\mu(x) - \nu(x))^2}{\mu(x) + \nu(x)} \right)} = \sqrt{2} \Delta(\mu, \nu) \end{aligned} \quad (4)$$

This quantity has an advantage compared to the total variation: for any $f : G \rightarrow \mathbb{R}$, we have (using Cauchy-Schwarz similarly)

$$|\mu(f) - \nu(f)| \leq \Delta(\mu, \nu) (\mu(f^2) + \nu(f^2))^{1/2}. \quad (5)$$

We will always be interested in the particular case of random walks. In order to lighten the notations, we set

$$\begin{aligned} \Delta_n(x, y) &:= \Delta(\mathcal{L}(X_n | X_0 = x), \mathcal{L}(X_{n-1} | X_0 = y)) \\ &= \Delta(\mathcal{L}(X_n | X_0 = x), \mathcal{L}(X_n | X_1 = y)), \end{aligned} \quad (6)$$

the last equality following by the Markov property. Recall that $\mathcal{L}(X | \mathcal{E})$ denotes the law of X conditioned on \mathcal{E} . Note that the second measure is the law of the random walk after $n - 1$ steps, so the definition is not symmetric in x and y .

Theorem 7. *Let (P, V, ρ) be a stationary environment. For every $n > 0$, we have*

$$\mathbb{E}(\Delta_n(\rho, X_1)^2) \leq 2(H_n - H_{n-1}) \quad (7)$$

(as usual \mathbb{E} is over both the environment and the randomness of X_1).

Before proving Theorem 7, we state a result from [9] concerning H_1^n . We isolate it from the rest of the proof because it is the only place where stationarity is used (stationarity replaces transitivity as used in the context of groups).

Lemma 8. *Let (P, V, ρ) be a stationary environment. For every $n > 0$, we have $H_1^n = H_{n-1} + H_1$.*

Proof. Fix $n > 0$. We first use the conditional entropy formula $H(X, Y) = H(X|Y) + H(Y)$. In our context this is simply

$$\begin{aligned} \mathbf{H}_1^n(P, V, \rho) &= \sum_{x, y \in G} \phi(\mathbf{P}_\rho(X_1 = x, X_n = y)) \\ &= \sum_{x \in G} \mathbf{P}_\rho(X_1 = x) \sum_{y \in G} \phi(\mathbf{P}_\rho(X_n = y | X_1 = x)) + \sum_{x \in G} \phi(\mathbf{P}_\rho(X_1 = x)) \end{aligned}$$

which we simplify using the Markov property giving

$$\mathbf{P}_\rho(X_n = y | X_1 = x) = \mathbf{P}_x(X_{n-1} = y).$$

Taking the expectation with respect to the environment we obtain

$$\begin{aligned} H_1^n &= \mathbb{E} \left[\sum_{x \sim \rho} \mathbf{P}_\rho(X_1 = x) \sum_{y \in G} \phi(\mathbf{P}_x(X_{n-1} = y)) \right] + \mathbb{E} \left[\sum_{x \in G} \phi(\mathbf{P}_\rho(X_1 = x)) \right] \\ &= \mathbb{E} [\mathbf{H}_{n-1}(P, V, X_1)] + \mathbb{E} [\mathbf{H}_1(P, V, \rho)] = H_{n-1} + H_1 \end{aligned}$$

where in the last equality we used the fact that (P, V, X_1) has the same law as (P, V, ρ) — formally, we use the coupling that makes them isomorphic with probability 1, and since a Markov chain isomorphism preserves the entropy, the expected entropy of the two environments (P, V, X_1) and (P, V, ρ) must be equal (this is not a property of entropy, it would be true for any function of the environment). \square

Before continuing let us state one corollary of Lemma 8 which is not necessary for the proof of Theorem 7 but does shed some light on the quantities involved.

Corollary 9. $H_n - H_{n-1}$ is decreasing.

Proof. By Lemma 8,

$$H_n - H_{n-1} = H_n - H_1^n + H_1 = \mathbb{E}[\mathbf{H}_n - \mathbf{H}_1^n] + H_1.$$

and $\mathbf{H}_n - \mathbf{H}_1^n$ can be written as the conditioned entropy $-H(X_1|X_n)$ where X_n is the random walk at time n (this statement is quenched). This, however, increases since

$$H(X_1|X_n) = H(X_1|X_n, X_{n+1}) \leq H(X_1|X_{n+1}) \quad (8)$$

where the equality is due to the fact that conditioned on X_n , knowing X_{n+1} gives you no information about what happened before time n . The inequality in (8) is a generic fact about entropy — conditioning on more information reduces the relative entropy. Hence $\mathbf{H}_n - \mathbf{H}_1^n$ decreases, and so does its expectation. \square

Returning to the proof of Theorem 7, let us recall another well-known fact. For any two random variables X and Y , $H(X, Y) \leq H(X) + H(Y)$ with equality holding if and only if X and Y are independent. The next lemma gives a quantitative version of this fact.

Lemma 10. *For any two random variables X and Y ,*

$$\sum_y \mathbb{P}(Y = y) \Delta^2(\mathcal{L}(X), \mathcal{L}(X|Y = y)) \leq 2(H(X) + H(Y) - H(X, Y)). \quad (9)$$

Proof. We start with the fact that for $t > 0$

$$2t \log t \geq \frac{(t-1)^2}{t+1} + 2t - 2 \quad (10)$$

(This can be seen by Taylor expanding $t \log t$ to the second order at 1.) For shortness denote

$$p(x) = \mathbb{P}(X = x) \quad p(y) = \mathbb{P}(Y = y) \quad p(x, y) = \mathbb{P}(X = x, Y = y).$$

Then the left-hand side of (9) is (recall the definition (3) of Δ),

$$\begin{aligned} \text{LHS} &= \sum_y p(y) \sum_x \frac{(p(x, y)/p(y) - p(x))^2}{p(x, y)/p(y) + p(x)} = \\ &= \sum_{y,x} p(x)p(y) \left(\frac{(p(x, y)/p(x)p(y) - 1)^2}{p(x, y)/p(x)p(y) + 1} + \underbrace{2 \frac{p(x, y)}{p(x)p(y)} - 2}_0 \right) \end{aligned}$$

where we were allowed to add the expression denoted by $\underbrace{}_0$ since summing over x and y makes both terms equal to 2 and cancel out. Thus, we get

$$\begin{aligned} \text{LHS} &\stackrel{(10)}{\leq} 2 \sum_{x,y} p(x)p(y) \left(\frac{p(x, y)}{p(x)p(y)} \log \frac{p(x, y)}{p(x)p(y)} \right) = \\ &= 2 \sum_{x,y} p(x, y) (\log p(x, y) - \log p(x) - \log p(y)) = 2(-H(X, Y) + H(X) + H(Y)) \end{aligned}$$

where in the last equality we used that $\sum_y p(x, y) = p(x)$ and $\sum_x p(x, y) = p(y)$. \square

Proof of Theorem 7. This is a direct corollary of Lemmas 8 and 10. Indeed, by Lemma 10,

$$\begin{aligned} \mathbf{E}(\Delta_n(\rho, X_1)^2) &= \sum_x \mathbf{P}(X_1 = x) \Delta(\mathcal{L}(X_n), \mathcal{L}(X_n|X_1 = x))^2 \\ &\leq 2(\mathbf{H}_1 + \mathbf{H}_n - \mathbf{H}_1^n). \end{aligned}$$

We now take expectation with respect to the environment and get from Lemma 8

$$\mathbb{E}(\Delta_n(\rho, X_1)^2) \leq 2\mathbb{E}(\mathbf{H}_1 + \mathbf{H}_n - \mathbf{H}_1^n) = 2(H_n - H_{n-1}). \quad \square$$

We are now in a position to prove Theorem 3:

Proof of Theorem 3. We only need to prove that for almost every environment, $h(\rho) = h(X_1)$ for any sublinear harmonic function. Indeed, stationarity would then imply that for almost every P , $h(X_n) = h(X_{n+1})$ for any sublinear harmonic function. Since the

Markov chain is irreducible, (X_n) can visit any vertex and we deduce that almost surely any sublinear harmonic function is constant.

For any harmonic function h with respect to the environment we have

$$h(\rho) = \mathbf{E}_\rho(h(X_n)) \quad h(X_1) = \mathbf{E}_{X_1}(h(X_{n-1})).$$

Hence

$$\begin{aligned} |h(\rho) - h(X_1)| &= |\mathbf{E}_\rho[h(X_n)] - \mathbf{E}_{X_1}[h(X_{n-1})]| \\ \text{By (5)} \quad &\leq \Delta_n(\rho, X_1) \sqrt{\mathbf{E}_\rho[h^2(X_n)] + \mathbf{E}_{X_1}[h^2(X_{n-1})]}. \end{aligned}$$

Taking expectation over X_1 gives

$$\begin{aligned} \mathbf{E}_\rho|h(\rho) - h(X_1)| &\leq \mathbf{E}_\rho\left[\Delta_n(\rho, X_1) \sqrt{\mathbf{E}_\rho[h^2(X_n)] + \mathbf{E}_{X_1}[h^2(X_{n-1})]}\right] \\ \text{by Cauchy-Schwarz} \quad &\leq \sqrt{2\mathbf{E}_\rho[\Delta_n(\rho, X_1)^2] \mathbf{E}_\rho[h^2(X_n)]} \end{aligned} \quad (11)$$

where in the last line we also used that $\mathbf{E}_\rho[\mathbf{E}_{X_1}[h^2(X_{n-1})]] = \mathbf{E}_\rho[h^2(X_n)]$.

Now, assume that h has sublinear growth. First, for any $\varepsilon > 0$, there exists a constant K such that for all $x \in V$,

$$h^2(x) \leq \varepsilon \mathbf{d}(x, \rho)^2 + K. \quad (12)$$

Second, the Markov chain has annealed polynomial growth, therefore the entropy satisfies

$$H_n \leq \mathbb{E}[\log |\mathbf{B}_\rho(n)|] \leq \log \mathbb{E}[|\mathbf{B}_\rho(n)|] \leq \log[Cn^d]$$

and is at most logarithmic (we use the fact that \log is concave). Hence $H_n - H_{n-1} \leq c/n$ for infinitely many n . Using Theorem 7 and (SDB) we get

$$\mathbb{E}[n\Delta_n(\rho, X_1)^2] + \mathbb{E}[n^{-1}\mathbf{d}(X_n, \rho)^2] \leq C \quad \text{for infinitely many } n.$$

Hence, by Fatou's Lemma, for almost every environment there exists $\mathbf{c}_1 < \infty$ such that

$$\mathbf{E}_\rho[n\Delta_n(\rho, X_1)^2] + \mathbf{E}_\rho[n^{-1}\mathbf{d}(X_n, \rho)^2] \leq \mathbf{c}_1 \quad \text{for infinitely many } n, \quad (13)$$

where this time the sequence of n for which it holds depends on the environment, *i.e.* is random. Putting (12) and (13) in (11), we deduce that for almost every environment, and for every h harmonic and sublinear on it,

$$\mathbf{E}_\rho(|h(\rho) - h(X_1)|) \leq \mathbf{c}_2 \varepsilon^{1/2}.$$

Letting ε go to 0, we deduce that $h(\rho) = h(X_1)$ almost surely for any sublinear harmonic function. \square

Inequality (11) relates the entropy to the value of possible harmonic functions at X_n . Its use is not restricted to the case of diffusive environments with polynomial growth. For instance, one can use this inequality to prove a characterization of almost sure Liouville property for stationary random graphs (this was proved in [9] using a more direct generalization of [46]). For completeness, we state the result in [9] here.

Corollary 11. [9] *Let (P, V, ρ) be a stationary environment. If H_n/n converges to 0, then P has the Liouville property (i.e. has no non-constant bounded harmonic functions) almost surely.*

We would like to emphasize why we use $\Delta(\mu, \nu)$. Csiszár's inequality [24, 25] relates the total variation between two measures to their relative entropy. In our context, an inequality involving the total variation can also be found, hence giving a bound on the best coupling (in time) between two random walks starting at neighbors. For completeness, we state the inequality here (it is a consequence of (4) applied to (7)): for a stationary environment (P, V, ρ) and $n > 0$, we have

$$\mathbb{E}(\|\mathcal{L}(X_n) - \mathcal{L}(X_n|X_1)\|_{TV}^2) \leq 4(H_n - H_{n-1}).$$

Interestingly, this inequality is not strong enough for our applications, since controlling the probability that two random walks merge before time n says nothing about their behavior when they do not couple.

Examples. We finish this section by presenting a collection of examples.

Example 2.1 (Random conductance). Consider the graph \mathbb{Z}^d and let ν be given by a shift-invariant law (for example i.i.d. positive random variables). We assume that the set of sites connected by edges with positive conductances is infinite. The random walk induces a Markov process on the environment, cf. Kipnis and Varadhan [50], called the environment as seen from the particle. This process can be made stationary by weighting each configuration proportionally to $\nu(\rho)$.

This model has been studied extensively. Under the assumption of *uniform ellipticity*: $\exists \alpha > 0 : \mathbb{P}[\alpha < \nu(x, y) < 1/\alpha] = 1$, many things are known on the environment. First, the Poincaré inequality is a direct consequence of the \mathbb{Z}^d case. Second, Delmotte proved in [28] that the Poincaré inequality implies that there exist $c_1, c_2 > 0$ such that

$$\mathbf{P}_\rho[X_n = x] < \frac{c_1}{n^{d/2}} e^{-c_2|x|^2/t}$$

(a corresponding lower bound also holds but is not needed for our purposes). Third, an annealed invariance principle holds in the sense that the law of the paths under the measure integrated over the environment scales to a non-degenerate Brownian motion [50]. In particular, Theorem 3 applies in this case.

Once the assumption of uniform ellipticity is relaxed, matters get more complicated. An example of random conductance models without uniform ellipticity is the infinite cluster of percolation which we will discuss next. For an unusual example of a transitive conductance model, see the work of Disertori, Spencer and Zirnbauer [32] who reduced a supersymmetric hyperbolic sigma model to the study of random walk on a certain (highly correlated) random environment.

Example 2.2 (Infinite cluster of percolation). Consider the percolation measure with a parameter p such that there exists an infinite cluster with probability 1. See [39] for details about percolation. Set \mathbb{P}_0 to be the law of the infinite cluster conditioned to contain 0.

As in the previous example, the random walk on ω induces a Markov chain on the space Ω of infinite subgraphs of \mathbb{Z}^d containing the origin. When weighting each configuration proportionally to the number of neighbors of the origin we obtain a stationary measure with respect to the shift along the random walk.

Since the infinite cluster of percolation can be seen as a stationary random graph with polynomial volume growth and since the random walk is diffusive [49, 4], Theorem 3 applies, and we get Theorem 1.

Example 2.3 (Centered random environments). This is our first non-reversible example. A centered random environment is, roughly speaking, a Markov chain on \mathbb{Z}^d such that the probabilities can be “decomposed” into a sum over cycles. Such environments, even when non-reversible, are still heuristically quite close to reversible, and in particular they have a stationary version which is related to the usual version by an explicit reweighting, like in the reversible case [31, Section 3]. See Deuschel and Kösters [31] for a proof of a CLT, which implies (SDB) — of course, a CLT is much stronger than (SDB) — and hence our results apply.

Example 2.4 (Balanced random environments). This is another non-reversible example, which is “farther” from reversible than the previous one. A balanced random environment is a Markov chain P with state space \mathbb{Z}^d and nearest neighbor movements, such that for every $x \in \mathbb{Z}^d$ and every unit vector e_i , $P(x, x + e_i) = P(x, x - e_i)$. It follows that X_n is a martingale, and hence (SDB) is an immediate corollary of the Azuma-Hoeffding inequality. The issue is therefore only stationarity. In the case that the environment μ is uniformly elliptic and stationary and ergodic *to the action of \mathbb{Z}^d* (this is different from our notion of stationarity!), Lawler [54, Theorem 3] showed that there exists a stationary measure (in our sense) λ which is mutually absolutely continuous with respect to μ . Hence our results apply to λ , and hence also to μ . Guo and Zeitouni weakened the requirement of uniform ellipticity to just ellipticity, at the price of restricting the environment to the i.i.d. case [40]. Berger and Deuschel [15] have removed the requirement of ellipticity altogether.

Example 2.5 (Random environments with cut points). Under certain conditions, one can prove that a random walk in non-reversible random environments in \mathbb{Z}^d , d large enough, has cut points, and deduce from that a CLT and the existence of a stationary environment, hence our techniques apply. See [17] for the details.

Let us give one example which is not embedded in \mathbb{Z}^d , and in fact has unbounded degrees.

Example 2.6 (Poisson point process). Examine a Poisson point process in \mathbb{R}^d . Add the point 0 (this is often called “the Palm process”) and let it be the root. Construct a graph by some process invariant under translations of \mathbb{R}^d . For example, connect any two points by an edge with weight which depends on their Euclidean distance [18] or construct the Delauney triangulation [37]. Give each configuration a “probability proportional to the total weight of 0”. The resulting process is stationary and diffusive (see *e.g.* [18, §2.1] or [37, Lemma A.1] for stationarity — subdiffusivity can be deduced from [8, Section 2] or from the two previous papers). Hence our theorem applies.

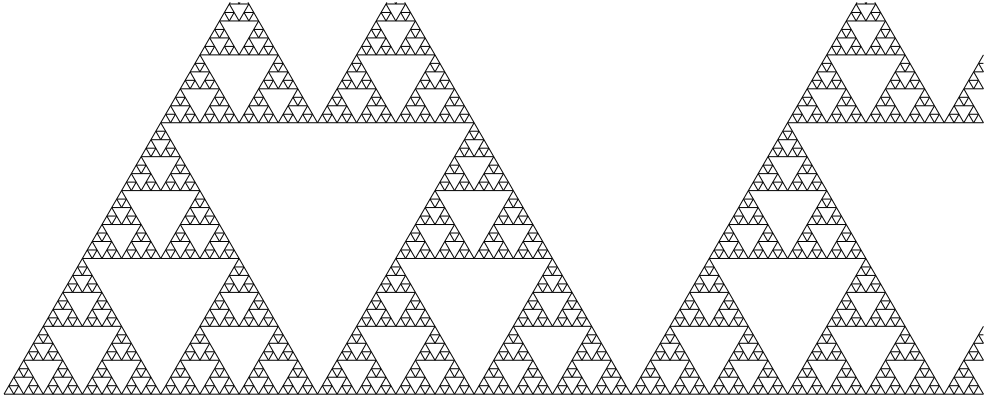


Figure 1: A portion of the graphical Sierpinski gasket.

The previous examples dealt with random walks which are diffusive. An interesting situation, which cannot hold in the case of groups, is environments with subdiffusive behavior. We give four examples of these.

Example 2.7 (Graphical fractals). A graphical fractal is a graph which is constructed like one of the classical fractals (the Sierpinski gasket, for example) but inside out — bigger pieces of the graph are constructed from smaller pieces by connecting them in a repeated fashion (see [5] for precise definitions and main properties). See figure 1 for an example, the graphical Sierpinski gasket. A graphical fractal always has an invariant measure and is always diffusive or subdiffusive, and in many examples is in fact subdiffusive, see *e.g.* [3]. Let us remark that a significant part in the remarkable work of Barlow and Bass on the Sierpinski carpet [6] has to do with the construction of a coupling, so a tool (like the one described in this section) that gives easy proofs that couplings exist should be useful.

Example 2.8 (Critical Galton-Watson trees). The critical Galton-Watson tree with any offspring distribution conditioned to survive is stationary (see [47, 59, 58]) and subdiffusive. If the offspring distribution has finite variance, the diffusivity exponent $1/3$ was proved in [49].

Example 2.9 (Infinite Incipient Cluster). Consider the critical percolation on \mathbb{Z}^d conditioned on the fact that the origin is connected to infinity [48] — one cannot really condition on this event which has probability 0 (proved in $d = 2$ and high d and conjectured in the others), but the object can be defined properly using a limit process. For example, one may take $p_c + \epsilon$ percolation, condition on $\vec{0}$ being in the cluster and then take a limit of the resulting measures as $\epsilon \rightarrow 0$. Since for each ϵ the measure is stationary (as usual after reweighting the configurations proportionally to the degree of $\vec{0}$), so will be their limit if it exists (or any subsequence limit in general). The limit is known to exist in 2 dimensions [48, 44] and in high dimensions [43, 42]. It was proved in [49, 52] that the random walk is subdiffusive on this cluster (in high dimension the diffusivity exponent is $1/3$, as on the tree). Since it is embedded in \mathbb{Z}^d , it grows no faster than polynomially. Therefore, there are no linear growth harmonic functions on it.

Example 2.10 (UIPQ). Let G_n be fixed or random finite graphs. Take ρ_n to be a random vertex in G_n , selected according to the stationary measure on G_n . Then the limit of

(G_n, ρ_n) , if it exists, is called the graph limit [13]. This limit is always stationary [53, Section 1.3].

A particular case is when G_n is a planar quadrangulation with n faces, chosen uniformly. The graph limit is known as the uniform infinite planar quadrangulation. It was proved to be subdiffusive with diffusivity exponent bounded from above by $1/3$ in [10]. Thus, there are no linear growth harmonic functions in this case as well.

Other growth rates. Examining the proof of Theorem 3 the same argument can also be used with growth rates bigger than polynomial. A general statement would be

Theorem 3'. *Let (P, V, ρ) be a stationary environment. Then there are no non-constant harmonic functions h which satisfy*

$$\mathbf{E}_\rho[h(X_n)^2] \cdot (H_n - H_{n-1}) \rightarrow 0$$

even if this holds just along a subsequence of n .

In particular this holds for fixed transitive graphs, which is a version of a result of [34, Section 5].

A remark on connectivity. We assumed throughout that the environment (P, V, ρ) is irreducible, *i.e.* that for any $v, w \in V$ there is some n such that $P^n(v, w) > 0$. This assumption was only used once: we showed that a not-necessarily-irreducible stationary environment satisfies that every harmonic function h has $h(\rho) = h(X_1)$ almost surely, and concluded, using irreducibility, that h is constant. The assumption of irreducibility is of course necessary, as a disconnected graph always has bounded non-constant harmonic functions, namely functions which are constant on each component, but with different constants.

Nevertheless, in the non-reversible case, the assumption of irreducibility can be weakened slightly: we only need to assume that for every v and w there exist n, m and x such that $P^n(v, x) > 0$ and $P^m(w, x) > 0$. The proof is the same — since $h(\rho) = h(X_1)$ almost surely then this gives that $h(v) = h(x) = h(w)$ almost surely and h is constant. An example of a stationary graph satisfying this is as follows: take a 3-regular tree T , choose a height function ℓ (*i.e.* a function such that each vertex has one neighbor with ℓ bigger by one, and two neighbors with ℓ smaller by one), and orient all edges “up” *i.e.* in the direction of the larger ℓ . Of course, the random walk on the resulting graph is so degenerate it can hardly be called random, as each vertex has only one outgoing edge. But this is irrelevant at this point. This environment is not irreducible in the usual sense, but does satisfy the weaker assumption and hence our results apply (again, in this case it is simple to analyze the harmonic functions directly). Multiplying by \mathbb{Z} will give a slightly less trivial example.

3 Polynomial growth harmonic functions

In this section we prove Theorem 4. The proof boils down to the observation that macroscopic Poincaré inequality and volume growth estimates are sufficient. The strategy fol-

lows the lines of Shalom and Tao [66, 68], where a quantitative version of Gromov's celebrated theorem on groups of polynomial growth (any group of polynomial growth is virtually nilpotent) is proved. The proof is inspired by an elegant proof of this theorem due to Kleiner [51] harnessing spaces of harmonic functions with polynomial growth in a crucial way. We start with a very general inequality, called the reverse Poincaré inequality, which holds in any graph. For the sake of completeness, we prove it in our context.

Proposition 12 (Reverse Poincaré inequality). *For any weighted graph (G, ν) and any function $h : G \rightarrow \mathbb{R}$ harmonic on a ball $\mathbf{B}_x(2n)$,*

$$\sum_{(y,z) \in E(\mathbf{B}_x(n))} (h(z) - h(y))^2 \nu(y, z) \leq \frac{4}{n^2} \sum_{y \in \mathbf{B}_x(2n)} h(y)^2 \nu(y). \quad (14)$$

for every $x \in G$ and $n > 0$.

Proof. For this proof, we denote the quantity $f(x)$ by f_x . Let $h : \mathbf{B}_x(2n) \rightarrow \mathbb{R}$ be harmonic and let ϕ be a function such that $\phi_y = 1$ for $y \in \mathbf{B}_x(n)$, $\phi_y = 0$ for $y \notin \mathbf{B}_x(2n-1)$ and $|\phi_y - \phi_z| \leq 1/n$ for all $y \sim z$. For example,

$$\phi_y := \min \left(1, 2 - \frac{d(y, x)}{n} \right).$$

We have

$$\sum_{E(\mathbf{B}_x(n))} (h_y - h_z)^2 \nu(y, z) = \sum_{E(\mathbf{B}_x(n))} \frac{1}{2} (\phi_y^2 + \phi_z^2) (h_y - h_z)^2 \nu(y, z). \quad (15)$$

To make the calculation a little shorter we represent the sum on the right-hand side of (15) as a sum of $\frac{1}{2} \phi_y^2 (h_y - h_z)^2$ over *directed edges*. Denote by E^* the set of directed edges in $B_x(2n)$ i.e. both (y, z) and (z, y) appear in E^* and are different. For an edge $(y, z) \in E^*$, an easy computation shows that $\phi_y^2 (h_z - h_y)^2$ equals the lengthy yet straightforward quantity

$$(h_z \phi_z^2 - h_y \phi_y^2)(h_z - h_y) - h_z (\phi_z - \phi_y)^2 (h_z - h_y) - 2h_z \phi_y (\phi_z - \phi_y)(h_z - h_y).$$

We start by dealing with the first term. Integration by parts and the fact that $h\phi^2$ vanishes on the boundary of $\mathbf{B}_x(2n)$ imply

$$\sum_{E^*} (h_z \phi_z^2 - h_y \phi_y^2)(h_z - h_y) \nu(y, z) = 2 \sum_{y \in \mathbf{B}_x(2n)} h_y \phi_y^2 \left(\sum_{z \sim y} (h_y - h_z) \nu(z, y) \right)$$

Since h is harmonic, this sum equals 0.

For the second term, since $|h_z(h_z - h_y)| \leq \frac{3}{2}h_z^2 + \frac{1}{2}h_y^2$ and $|\phi_z - \phi_y| \leq 1/n$, we have that each summand is bounded by $(3h_z^2 + h_y^2)/(2n^2)$. When summing over E^* we obtain

$$\left| \sum_{E^*} h_z (\phi_z - \phi_y)^2 (h_z - h_y) \nu(y, z) \right| \leq \frac{2}{n^2} \sum_{\mathbf{B}_x(2n)} h_y^2 \nu(y).$$

For the third term, note that

$$|h_z \phi_y(\phi_z - \phi_y)(h_z - h_y)| \leq \frac{1}{4}(h_y - h_z)^2 \phi_y^2 + h_z^2(\phi_z - \phi_y)^2. \quad (16)$$

So,

$$\begin{aligned} & \sum_{E^*} |h_z \phi_y(\phi_z - \phi_y)(h_z - h_y)| \nu(z, y) \\ \text{By (16)} \quad & \leq \frac{1}{4} \sum_{E^*} (h_y - h_z)^2 \phi_y^2 \nu(y, z) + \sum_{E^*} h_z^2 (\phi_z - \phi_y)^2 \nu(y, z) \\ & \leq \frac{1}{4} \sum_{E^*} (h_y - h_z)^2 \phi_y^2 \nu(y, z) + \frac{1}{n^2} \sum_{\mathbf{B}_x(2n)} h_y^2 \nu(y) \end{aligned}$$

using the bound $|\phi_z - \phi_y| \leq \frac{1}{n}$ for every $y \sim z$. Putting the bound on the different terms together leads to

$$\sum_{E^*} (h_y - h_z)^2 \phi_y^2 \nu(y, z) \leq \frac{1}{2} \sum_{E^*} (h_y - h_z)^2 \phi_y^2 \nu(y, z) + \frac{4}{n^2} \sum_{\mathbf{B}_x(2n)} h_y^2 \nu(y).$$

which gives

$$\sum_{E(\mathbf{B}_x(n))} (h_z - h_y)^2 d\nu(y, z) \leq \frac{1}{2} \sum_{E^*} (h_z - h_y)^2 \phi_y^2 \nu(y, z) \leq \frac{4}{n^2} \sum_{\mathbf{B}_x(2n)} h_y^2 \nu(y). \quad \square$$

Lemma 13. *Let (G, ν) be a random graph satisfying the volume doubling condition $(VD)_G$ almost surely. Then there exists $\mathbf{c} > 0$ such that for any finite λ and n large enough, there is a covering of the ball $\mathbf{B}_\rho(\lambda n)$ by less than M_λ balls $\mathbf{B}_{y_1}(n), \dots, \mathbf{B}_{y_k}(n)$ satisfying that every point $x \in \mathbf{B}_\rho(n)$ belongs to at most \mathbf{c} balls $\mathbf{B}_{y_i}(2n)$.*

Further, \mathbf{c} depends only on the volume doubling constant \mathbf{C}_{VD} , and M_λ depends only on λ and \mathbf{C}_{VD} .

We call a covering with this property *proper*.

Proof. Let λ and G be as above. Let n be large enough so that $(VD)_G$ holds for our λ . Given this, we can choose a maximal family of disjoint balls $\mathbf{B}_{y_1}(n/2), \dots, \mathbf{B}_{y_k}(n/2)$ with $y_j \in \mathbf{B}_\rho(\lambda n)$ for all j .

- Since the family $\{\mathbf{B}_{y_j}(n/2)\}$ is maximal, every vertex in $\mathbf{B}_\rho(\lambda n)$ must be within distance $\leq n$ from one of the y_j , so $\mathbf{B}_\rho(\lambda n)$ is covered by $\mathbf{B}_{y_1}(n), \dots, \mathbf{B}_{y_k}(n)$.
- For any $x \in \mathbf{B}_\rho(\lambda n)$, if $x \in \mathbf{B}_{y_j}(2n)$, then $\mathbf{B}_{y_j}(n/2) \subset \mathbf{B}_x(3n)$. Using volume doubling we see that $\nu(\mathbf{B}_x(3n)) \leq \mathbf{C}_{\text{VD}}^4 \nu(\mathbf{B}_{y_j}(n/2))$, hence (since these balls are disjoint) we have that the number of y_j such that $x \in \mathbf{B}_{y_j}(2n)$ is at most \mathbf{C}_{VD}^4 .
- We use volume doubling similarly and get

$$\nu(\mathbf{B}_\rho((\lambda + 1)n)) \leq \mathbf{C} \nu(\mathbf{B}_{y_j}(n/2)) \quad \forall j.$$

(the constant is $\mathbf{C}_{\text{VD}}^{\lceil \log_2(\lambda+1) \rceil + 2}$). Since these balls are all disjoint and fully contained in $\mathbf{B}_\rho((\lambda+1)n)$, we get

$$k \min_j \nu(\mathbf{B}_{y_j}(n/2)) \leq \nu\left(\bigcup_j \mathbf{B}_{y_j}(n/2)\right) \leq \nu(\mathbf{B}_\rho((\lambda+1)n)) \leq \mathbf{C} \min_j \nu(\mathbf{B}_{y_j}(n/2))$$

and we get that the number of balls k is bounded by the same \mathbf{C} . \square

Lemma 14. *Let (G, ν) be a random graph satisfying $(P)_G$ almost surely. For almost every G , there exists $\mathbf{c} > 0$ such that for every $\varepsilon > 0$ and n large enough, and for every proper covering of $\mathbf{B}_\rho(n)$ by balls of radius εn , if $h : G \rightarrow \mathbb{R}$ is harmonic and has 0 mean on all the balls of the covering then*

$$\sum_{\mathbf{B}_\rho(n)} h(z)^2 \nu(z) \leq \mathbf{c} \varepsilon^2 \sum_{\mathbf{B}_\rho(4n)} h(z)^2 \nu(z). \quad (17)$$

Further, \mathbf{c} depends only on \mathbf{C}_P , the constant in the Poincaré inequality.

Proof. Let G be an environment and fix n large enough so that $(P)_G$ holds true for $\lambda = 1/\varepsilon$ and εn : for every $x \in \mathbf{B}_\rho(n)$ and f a map on $\mathbf{B}_\rho(n)$,

$$\sum_{y \in \mathbf{B}_x(\varepsilon n)} (f(y) - \bar{f}_{\mathbf{B}_x(\varepsilon n)})^2 \nu(y) \leq \mathbf{C}_P(\varepsilon n)^2 \sum_{(y,z) \in E(\mathbf{B}_x(2\varepsilon n))} |f(y) - f(z)|^2 \nu(y, z).$$

Consider $h : G \rightarrow \mathbb{R}$ a harmonic function and $\mathbf{B}_{y_1}(\varepsilon n), \dots, \mathbf{B}_{y_k}(\varepsilon n)$ the proper covering of $\mathbf{B}_\rho(n)$ from the statement of the lemma. The hypothesis asserts that $\bar{h}_{\mathbf{B}_{y_i}(\varepsilon n)} = 0$ for every i , so that Poincaré inequality implies

$$\begin{aligned} \sum_{\mathbf{B}_{y_i}(\varepsilon n)} h^2(z) \nu(z) &= \sum_{\mathbf{B}_{y_i}(\varepsilon n)} (h(z) - \bar{h}_{\mathbf{B}_{y_i}(\varepsilon n)})^2 \nu(z) \\ &\leq \mathbf{C}_P \varepsilon^2 n^2 \sum_{E(\mathbf{B}_{y_i}(2\varepsilon n))} (h(z) - h(t))^2 \nu(z, t). \end{aligned}$$

Since the $\mathbf{B}_{y_i}(\varepsilon n)$ have uniformly bounded overlap (each point belong to at most \mathbf{c} balls), and since $\mathbf{B}_{y_i}(2\varepsilon n) \subset \mathbf{B}_\rho(2n)$, we find

$$\sum_{\mathbf{B}_\rho(n)} h^2(z) \nu(z) \leq \mathbf{c} \mathbf{C}_P \varepsilon^2 n^2 \sum_{E(\mathbf{B}_\rho(2n))} (h(z) - h(t))^2 \nu(z, t). \quad (18)$$

Using the reverse Poincaré inequality (Proposition 12) for the larger ball, we conclude

$$\sum_{\mathbf{B}_\rho(n)} h^2(z) \nu(z) \leq 4\mathbf{c} \mathbf{C}_P \varepsilon^2 \sum_{\mathbf{B}_\rho(4n)} h^2(z) \nu(z) \quad (19)$$

which implies the claim with $\mathbf{c} = 4\mathbf{c} \mathbf{C}_P$. \square

Proof of Theorem 4. We aim to prove that the space of harmonic functions u such that $|u(x)| \leq \mathbf{C} d(\rho, x)^k$ for every $x \in G$ is finite dimensional. Consider an environment G satisfying $(VD)_G$ and $(P)_G$. Let \mathbf{c}_2 and \mathbf{c} be large enough so that the two previous

lemmas hold true. On the set of harmonic functions on $\mathbf{B}_\rho(n)$, a scalar product between two functions can be defined by

$$\langle f, g \rangle_n = \sum_{\mathbf{B}_\rho(n)} f(x)g(x)\nu(x).$$

Consider d harmonic functions u_1, \dots, u_d on G and set $V = \text{span}(u_1, \dots, u_d)$.

Let $\varepsilon > 0$ be some parameter to be fixed later. For n large enough, there exists a proper covering $\mathbf{B}_{y_1}(\varepsilon n), \dots, \mathbf{B}_{y_M}(\varepsilon n)$ of $\mathbf{B}_\rho(n)$ by $M = M_{1/\varepsilon}$ balls. Therefore there is a codimension $d - M$ vector space $V_0 \subset V$ of harmonic functions with mean 0 on each of the balls $\mathbf{B}_{y_i}(\varepsilon n)$. Let v_1, \dots, v_d be an orthogonal basis of V for $\langle \cdot, \cdot \rangle_{4n}$ such that v_1, \dots, v_{d-M} is a basis of V_0 . Examine the Gram matrix of v_i i.e. the $d \times d$ matrix whose entries are $\langle v_i, v_j \rangle_n$. Then

$$\begin{aligned} \det [\langle v_i, v_j \rangle_n]_{i,j} &\leq \prod_{i=1}^d \langle v_i, v_i \rangle_n \\ &\leq \prod_{i=1}^{d-M} \mathbf{c}\varepsilon \langle v_i, v_i \rangle_{4n} \prod_{i=d-M+1}^d \langle v_i, v_i \rangle_{4n} \\ &= (\mathbf{c}\varepsilon)^{d-M} \det [\langle v_i, v_j \rangle_{4n}]_{i,j}, \end{aligned}$$

where in the first line we have used Hadamard's inequality, in the second Lemma 14 and in the last, the fact that (v_i) is orthogonal for $\langle \cdot, \cdot \rangle_{4n}$. Now, the ratio of two Gram determinants is preserved by linear operations on vectors, so we can return from the basis $\{v_i\}$ (which was specific to n) to our "original" basis $\{u_i\}$. We get

$$\det [\langle u_i, u_j \rangle_n]_{i,j} \leq (\mathbf{c}\varepsilon)^{d-M} \det [\langle u_i, u_j \rangle_{4n}]_{i,j}.$$

Iterating the reasoning, we find for every $r > 0$

$$\det [\langle u_i, u_j \rangle_n]_{i,j} \leq [(\mathbf{c}\varepsilon)^{d-M}]^r \det [\langle u_i, u_j \rangle_{4^r n}]_{i,j}.$$

Since every entry of the matrix is smaller than $\mathbf{C}(4^r n)^k$ thanks to the bound on the growth of the harmonic functions, we find

$$\det [\langle u_i, u_j \rangle_n]_{i,j} \leq d! n^{kd} \mathbf{C}^d ((\mathbf{c}\varepsilon)^{d-M} 4^{kd})^r.$$

We now fix ε to be $4^{-2k}/\mathbf{c}$. If $d > 2M$ then this would imply

$$(\mathbf{c}\varepsilon)^{d-M} 4^{kd} = 4^{-2k(d-M)+kd} < 1$$

and the right hand side would converges to 0. We deduce that $\det [\langle u_i, u_j \rangle_n]_{i,j} = 0$, and that the u_i restricted to the ball of radius n form a dependent family. Since this is true for every n large enough, we easily deduce that (u_i) is a linearly dependent family. The result holds for any family of d harmonic functions with growth bounded by $\mathbf{C}d(\cdot, \rho)^k$. It implies that the dimension of the vector space of harmonic functions with such growth is smaller or equal to $2M$. \square

Example 3.1 (Infinite cluster of percolation). The infinite cluster of percolation satisfies $(VD)_\omega$ and $(P)_\omega$ almost surely [4]. Therefore, spaces of harmonic functions with prescribed polynomial growth are finite dimensional.

Example 3.2 (Random conductance). Random conductances with uniform elliptic conditions also satisfy $(VD)_\omega$ and $(P)_\omega$ deterministically. Therefore, spaces of harmonic functions with prescribed polynomial growth are finite dimensional.

Example 3.3 (Wedges). Let f be some *slowly varying* function from $[0, \infty) \rightarrow [0, \infty)$. Define the wedge with respect to d and f to be

$$W := \{x \in \mathbb{Z}^d : |x_d| \leq f(|x_1| + \dots + |x_{d-1}|)\}.$$

Then it is well-known and not difficult to see that W (with the graph structure inherited from \mathbb{Z}^d) satisfies volume doubling and Poincaré inequality. Under some weak conditions on f and d (which we will not detail here, as that would take us too off-topic) so would percolation on W . Hence both W and supercritical percolation on it have a finite dimensional space of harmonic functions.

4 Linearly growing harmonic functions on the infinite cluster of percolation

In this section, we fix $d > 0$ and $p > p_c(d)$. The infinite cluster of supercritical percolation can be thought of as an approximation of \mathbb{Z}^d . In particular, macroscopic properties of the cluster are the same as those of \mathbb{R}^d . For instance, the random walk satisfies an invariance principle $(CLT)_\omega$ [67, 14, 62]: define

$$\tilde{B}_n(t) := \frac{1}{\sqrt{n}}(X_{tn}),$$

where for non-integer tn we define X_{tn} as the linear interpolation between $X_{\lfloor tn \rfloor}$ and $X_{\lceil tn \rceil}$; that is, $X_{tn} = X_{\lfloor tn \rfloor}(tn - \lfloor tn \rfloor) + X_{\lceil tn \rceil}(\lceil tn \rceil - tn)$. There exists $\sigma(d)$ such that the law of $(\tilde{B}_n(t), 0 < t < \infty)$ converges weakly to the law of a Brownian motion with variance $\sigma(d)$ as $n \rightarrow \infty$. The main step in the proof in all three papers [67, 14, 62] is the construction of a d -dimensional space of harmonic functions $\{f_v\}_{v \in \mathbb{R}^d}$ such that f_v has slope v i.e. $f_v(x) = \langle v, x \rangle + o(|x|)$. Let us state this as a theorem.

Theorem 15 ([67, 14, 62]). *Let $d \geq 2$, and $p > p_c(d)$. Let ω be the infinite cluster of percolation on \mathbb{Z}^d with parameter p . Then, for almost every ω there exists $\chi : \omega \rightarrow \mathbb{R}^d$ such that $x \mapsto x + \chi(x)$ is harmonic on ω , and*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sup_{x \in \mathbf{B}_\rho(n)} |\chi(x)| = 0 \quad a.s. \quad (20)$$

This (random) function is called the corrector.

With the constant functions, we get a $(d+1)$ -dimensional space of harmonic functions with (sub-)linear growth. Our aim in this section is to prove Theorem 5 from the introduction, namely that there are no other harmonic functions of linear growth.

Proof outline. Let h be a harmonic function with linear growth. Define $h_n : \mathbb{R}^d \rightarrow \mathbb{R}$ such that $h_n(x) = h(nx)/n$. In order to prove Theorem 5, we first show that (h_n) forms a precompact family (one can say that h has a scaling limit). The second step is to identify the possible limits. For this, we use the average property at the discrete level and the invariance principle to prove that limits are harmonic on \mathbb{R}^d . If the space of limits is at most d -dimensional, one can then use the absence of non-constant sublinear harmonic functions to show that the space of harmonic functions with linear growth is $(d+1)$ -dimensional.

Properties of the supercritical cluster. Recall that the infinite supercritical cluster of percolation ω can be seen as a stationary random graph with polynomial growth. It is well-known that the system is ergodic with respect to the shift by X_1 , see *e.g.* [14, Theorem 3.1]. Typical balls have the same growth as in the ambient space \mathbb{Z}^d in the following sense: there exists constants c and C such that for any finite λ , any n sufficiently large and any $x \in \mathbf{B}_\rho(\lambda n)$

$$cn^d \leq \nu(\mathbf{B}_x(n)) \leq Cn^d. \quad (21)$$

Clearly, (21) implies volume doubling $(VD)_\omega$. Moreover, the graph satisfies $(P)_\omega$ almost surely. Both properties were proved by Barlow [4]. Actually, Barlow proved quantitatively stronger versions of (21) and $(P)_\omega$: he obtained the volume growth estimates and the Poincaré inequality for every ball of radius larger than $C \log n$ in $\mathbf{B}_\rho(n)$. These improved versions allow to prove Harnack inequalities and Gaussian estimates (1) on the heat kernel. In [4] these results are stated for continuous time random walk, but they hold also for simple random walk, as was explained in [7, Section 2]. We do not need the full force of Gaussian estimates here — in particular we do not need far off-diagonal lower bounds which are particularly difficult — so let us make a list of corollaries from these Gaussian estimates which we will use.

Corollary 16. *For every finite λ , every n sufficiently large and every $x \in \mathbf{B}_\rho(\lambda n)$,*

$$\mathbf{P}_x(X_{n^2} = y) \leq Cn^{-d} \exp[-C|x - y|^2/n] \quad \forall y. \quad (22)$$

This immediately implies

$$\mathbf{E}_x[|X_{n^2} - x|^2] \leq c_3 n^2 \quad (23)$$

for some constant c_3 depending on the environment.

The lower bound has some periodicity requirements, because clearly $\mathbf{P}_x(X_t = y) = 0$ whenever $t + \sum(x_i - y_i)$ is odd.

Corollary 17. *For every finite λ , every n sufficiently large and every $x \in B(\lambda n)$,*

$$\mathbf{P}_x(X_{n^2} = y) \geq Cn^{-d} \quad \forall y \in \mathbf{B}_x(n) \text{ such that } n^2 + \sum(x_i - y_i) \text{ is even} \quad (24)$$

We start with a technical lemma which allow to move estimates at ρ to estimates at most points. The main ingredient is ergodicity.

Lemma 18. *Let $f(x, y, \omega)$ be some positive translation-invariant variable i.e. $f(x + s, y + s, \omega + s) = f(x, y, \omega)$, with $M := \mathbb{E}f(0, X_1, \omega) < \infty$. Then for every $\lambda > 0$ and for almost every environment ω ,*

$$\sum_{(x, y) \in E(\mathbf{B}_a(n))} f(x, y, \omega) \nu(x, y) \leq C \cdot M \cdot n^d$$

for n sufficiently large, and for all $a \in \mathbf{B}_\rho(\lambda n)$.

Proof. We will not denote the dependency of f on ω from this point on. By ergodicity [14, Theorem 3.1], for almost every ω , $\frac{1}{k} \sum_{t=0}^{k-1} f(X_t, X_{t+1})$ converges in L^1 to M (L^1 here is with respect to the measure \mathbf{P}_ρ). Therefore

$$\frac{1}{k} \sum_{t=0}^{k-1} \mathbf{E}_\rho[f(X_t, X_{t+1})] \leq 2M \text{ for } k \text{ sufficiently large.} \quad (25)$$

Moreover, the heat kernel lower bound (24) implies that there exists $C > 0$ such that for every t sufficiently large and every $x \in \mathbf{B}_\rho(\sqrt{t})$ with $t + \sum x_i$ even,

$$\mathbf{P}_\rho[X_t = x] \geq \frac{C}{|\mathbf{B}_\rho(\sqrt{t})|}$$

which we sum over t to get, for $x \sim y \in \mathbf{B}_\rho(k)$ and k sufficiently large

$$\frac{1}{2k^2} \sum_{t=0}^{2k^2} \mathbf{P}_\rho[X_t = x, X_{t+1} = y] \geq \frac{C_4}{|\mathbf{B}_\rho(k)|}.$$

Hence,

$$\sum_{(x, y) \in E(\mathbf{B}_\rho(k))} f(x, y) \nu(x, y) \leq \frac{|\mathbf{B}_\rho(k)|}{C_4} \cdot \frac{1}{2k^2} \sum_{t=0}^{2k^2} \mathbf{E}_\rho[f(X_t, X_{t+1})] \stackrel{(25)}{\leq} \frac{2M}{C_4} |\mathbf{B}_\rho(k)| \quad (26)$$

for k sufficiently large. Fix some k_0 such that the probability that (26) holds for $k > k_0$ is bigger than $\frac{1}{2}$. Call a point $b \in \omega$ *good* if

$$\sum_{(x, y) \in E(\mathbf{B}_b(k))} f(x, y) \nu(x, y) \leq \frac{2M}{C_4} |\mathbf{B}_b(k)| \quad \forall k > k_0$$

We know that ρ is good with probability greater than $\frac{1}{2}$.

By ergodicity again, for almost every environment ω ,

$$\frac{1}{k} \sum_{t=0}^{k-1} \mathbf{1}_{\{X_t \text{ is good}\}} \longrightarrow \mathbb{P}[\rho \text{ is good}] \quad \mathbf{P}_\rho\text{-almost surely} \quad (27)$$

when k goes to infinity. Now, consider an environment ω and fix $\delta > 0$. The maximum distance between two good points in the box $\mathbf{B}_\rho(n)$ behaves like $o(n)$. Indeed, if it was not the case there would exist $\varepsilon > 0$ such that with positive \mathbf{P}_ρ -probability, the random

walk X would not visit a good point during an interval εn before time n for an infinite number of n (here we used implicitly the invariance principle to see that the walk has positive probability to reach that “bad ball”, and stay there). This would prevent the left hand side in (27) from converging \mathbf{P}_ρ -almost surely as during these intervals the sum $\sum \mathbf{1}_{\{X_t \text{ is good}\}}$ would not increase while k would increase by a factor of $1 + \varepsilon$, causing the average to “plunge”. This cannot happen infinitely many times for a converging sequence.

Therefore, for every $a \in \mathbf{B}_\rho(\lambda n)$, there exists a good point b at distance less than n . Thus,

$$\begin{aligned} \sum_{(x,y) \in E(\mathbf{B}_a(n))} f(x,y) \nu(x,y) &\leq \sum_{(x,y) \in E(\mathbf{B}_b(2n))} f(x,y) \nu(x,y) \\ &\leq \frac{2M}{C_4} |\mathbf{B}_b(2n)| \stackrel{(21)}{\leq} CMn^d. \end{aligned} \quad \square$$

Recall the h_n from the proof sketch on page 22. There we defined $h_n(x) = h(nx)/n$ which is *a priori* only defined on the contracted infinite cluster. For simplicity let us extend it to all \mathbb{R}^d , *e.g.* by extending h to \mathbb{Z}^d by taking the value at the closest point of the infinite cluster, and then to \mathbb{R}^d by dividing each square $(n,m) + [0,1]^2$ into two triangles and interpolating on the triangles linearly. Once h is extended to all \mathbb{R}^d , so is h_n .

Proposition 19. *For almost every environment ω , any harmonic function h on ω with linear growth satisfies that for every compact $K \subset \mathbb{R}^d$, the sequence $(h_n)|_K$ is uniformly bounded and equicontinuous.*

Proof. Fix a harmonic map h with (at most) linear growth on an environment ω . There exists $A > 0$ such that $|h(x)| \leq A|x|$. We only need to prove equicontinuity on the ball (other compact sets K work the same). To do so, we prove that for any $\eta > 0$, there exists $\delta > 0$ such that $(h(a) - h(b))^2 \leq \eta n^2$ for any two points $a, b \in \mathbf{B}_\rho(n)$ at distance δn of each other, when n is large enough. For this reason, we will always assume that n is large enough so that $(P)_\omega$ and the d -dimensional volume growth (21) hold true for an appropriate λ .

Let $\delta, \varepsilon > 0$ to be fixed later (think about $\varepsilon \ll \delta$) and $a, b \in \mathbf{B}_\rho(n)$ with $\mathbf{d}(a, b) \leq \delta n$. Let B be some ball of radius $2\delta n$ containing both $\mathbf{B}_a(\delta n)$ and $\mathbf{B}_b(\delta n)$ — for example around the middle point of $[ab]$. Let \bar{h} be the average $\frac{1}{\nu(B)} \sum_{x \in B} h(x) \nu(x)$. Since $|h(a) - h(b)| \leq |h(a) - \bar{h}| + |h(b) - \bar{h}|$, it is enough to estimate these terms, and we estimate $|h(a) - \bar{h}|$. The other term is symmetric.

Set \mathcal{E} to be the event that $|X_{(\varepsilon n)^2} - a| \geq \delta n$. Note that

$$\mathbf{P}_a(\mathcal{E}) \leq \frac{\mathbf{E}_a(|X_{(\varepsilon n)^2} - a|^2)}{(\delta n)^2} \stackrel{(23)}{\leq} \frac{\mathbf{c}_3(\varepsilon n)^2}{(\delta n)^2} = \mathbf{c}_3(\varepsilon/\delta)^2$$

where the Markov inequality was used in the first inequality and the quenched diffusive behavior (23) in the second.

Now, we have

$$\begin{aligned} |h(a) - \bar{h}|^2 &\leq \left(\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}|] \right)^2 \\ &\leq 2 \left(\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}}] \right)^2 + 2 \left(\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}^c}] \right)^2 \end{aligned}$$

We first deal with the first term on the right:

$$\begin{aligned} \left(\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}}] \right)^2 &\leq \mathbf{E}_a \left[\left(|h(X_{(\varepsilon n)^2})| + |\bar{h}| \right)^2 \right] \cdot \mathbf{P}_a(\mathcal{E}) \\ &\leq \left(2\mathbf{E}_a[h(X_{(\varepsilon n)^2})^2] + 2\bar{h}^2 \right) \cdot \mathbf{P}_a(\mathcal{E}) \\ \text{Since } h(x) &\leq A|x| \quad \leq 2A \left(\mathbf{E}_a[|X_{(\varepsilon n)^2}|^2] + (1 + \delta)^2 n^2 \right) \cdot \mathbf{P}_a(\mathcal{E}) \\ \text{by (23)} \quad &\leq 2A(\mathbf{c}_3(1 + \varepsilon^2) + (1 + \delta)^2)n^2 \cdot \mathbf{c}_3 \frac{\varepsilon^2}{\delta^2} = \mathbf{c}_5 n^2 \frac{\varepsilon^2}{\delta^2} \end{aligned}$$

where in the first inequality we have used Cauchy-Schwarz.

For the second term, the heat kernel upper bound (22) shows that $\mathbf{P}_a(X_{(\varepsilon n)^2} = x) \leq C_6/(\varepsilon n)^d$ for any $x \in \mathbf{B}_\rho(n)$ and n large enough. Therefore,

$$\begin{aligned} \left(\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}^c}] \right)^2 &\leq \mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}|^2 \mathbf{1}_{\mathcal{E}^c}] \\ &\leq \frac{C_6}{(\varepsilon n)^d} \sum_{x \in \mathbf{B}_a(\delta n)} |h(x) - \bar{h}|^2 \nu(x) \\ &\leq \frac{C_6}{(\varepsilon n)^d} \sum_{x \in B} |h(x) - \bar{h}|^2 \nu(x). \end{aligned}$$

Harnessing Poincaré inequality, we find

$$\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}^c}]^2 \leq \frac{\mathbf{C}_P C_6}{(\varepsilon n)^d} \delta^2 n^2 \sum_{(x,y) \in E(B')} |h(x) - h(y)|^2 \nu(x, y),$$

where B' is the ball with same center as B and radius $4\delta n$.

Now, the quantity Δ_n controls the gradient of an harmonic function. Indeed, the same reasoning as the one used to derive (11) implies that

$$\begin{aligned} |h(x) - h(y)|^2 &\leq (\mathbf{E}_x[|h(X_n)|^2] + \mathbf{E}_y[|h(X_n)|^2]) \Delta_n(x, y)^2 \\ &\leq A^2 (\mathbf{E}_x[|X_n|^2] + \mathbf{E}_y[|X_n|^2]) \Delta_n(x, y)^2 \end{aligned}$$

for every n . Using diffusivity and taking the \liminf , we obtain

$$|h(x) - h(y)|^2 \leq \mathbf{c}_7 \liminf_{n \rightarrow \infty} \sqrt{n} \Delta_n(x, y)^2$$

where \mathbf{c}_7 does not depend on the points x, y (though it does depend on h through A). Denote this \liminf by $\Delta_\infty(x, y)^2$. We get

$$\left(\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}^c}] \right)^2 \leq \frac{\delta^2 \mathbf{C}_P C_6}{\varepsilon^d n^{d-2}} \sum_{(x,y) \in E(B)} \mathbf{c}_7 \Delta_\infty(x, y)^2 \nu(x, y).$$

We next note that $\mathbb{E}\Delta_\infty(\rho, X_1)^2 < \infty$. Indeed, the infinite cluster of percolation is a subgraph of \mathbb{Z}^d , it has uniform polynomial growth and $H_n \leq C_1 \log n$ for every n . Theorem 7 implies that $\mathbb{E}[\Delta_n(\rho, X_1)^2] \leq C_2/n$ for an infinite number of n . Using Fatou's lemma, we obtain that $\mathbb{E}[\Delta_\infty(\rho, X_1)^2] < \infty$. Thus we may use Lemma 18 for the function $f = \Delta_\infty^2$ and get (with the fact that B' has radius $4\delta n$),

$$\mathbf{E}_a[|h(X_{(\varepsilon n)^2}) - \bar{h}| \mathbf{1}_{\mathcal{E}^c}]^2 \leq \mathbf{c}_8 \frac{\delta^{d+2}}{\varepsilon^d} n^2.$$

Putting together the estimates for the two terms, we obtain

$$(h(a) - h(b))^2 \leq n^2 \left(\mathbf{c}_3 \frac{\varepsilon^2}{\delta^2} + \mathbf{c}_8 \frac{\delta^{d+2}}{\varepsilon^d} \right)$$

which implies the claim provided $\delta = \varepsilon^{(d+1)/(d+2)}$. \square

A final tool for the proof of Theorem 5 is a slight strengthening of Theorem 3 for the case of percolation. Theorem 3 required the function to be sublinear at every scale, but here it will be important to relax this assumption to work at only an (arbitrarily sparse) sequence of scales.

Lemma 20. *For supercritical percolation, $H_n - H_{n-1} \leq C/n$ for every n , where C is a constant depending only on d and p .*

(again, in Section 2 we showed that it holds for a sequence of n , here we need the claim for all n).

Proof. The heat kernel estimates (1) show, after a little calculation, that

$$\mathbf{H}_n = \frac{d}{2} \log n + O(1) \quad \forall n > n_0(\omega). \quad (28)$$

For $n \leq n_0(\omega)$ we can use a much rougher bound, say $\mathbf{H}_n \leq d \log(2n)$ which follows from the fact that for any cluster ω the distribution of R_n is supported on the cube $\{-n, \dots, n\}^d$ and any measure has entropy smaller than the entropy of the uniform measure on its support. Since $n_0(\omega)$ has a stretched exponential tail, we can integrate over the environment and get that $H_n = \frac{d}{2} \log n + O(1)$. This means that $H_n - H_{n/2} \leq C$ for some C . Using the fact that $H_n - H_{n-1}$ is decreasing (Corollary 9 on page 10) proves the lemma. \square

Corollary 21. *With probability 1, every harmonic function h on a supercritical percolation cluster which satisfies that for some $n_k \rightarrow \infty$,*

$$\mathbf{E}[h(X_{n_k})^2] = o(n_k)$$

is constant.

Proof. The proof is identical to that of Theorem 3. Equation (11) still holds for every n :

$$\mathbf{E}_\rho |h(\rho) - h(X_1)| \leq \sqrt{2\mathbf{E}_\rho [\Delta_n(\rho, X_1)^2] \mathbf{E}_\rho [h^2(X_n)]}.$$

We know from the previous lemma that $H_n - H_{n-1} \leq C/n$ for all n , and in particular for the n_k for which we have $\mathbf{E}[h(X_{n_k})^2] = o(n_k)$. By Theorem 7 this gives

$$\mathbb{E}[\Delta_{n_k}(\rho, X_1)^2] \leq C/n_k.$$

Using Fatou's lemma we know that there is some random \mathbf{c} and a random subsequence n'_k of n_k for which

$$\mathbf{E}[\Delta_{n'_k}(\rho, X_1)^2] \leq \mathbf{c}/n'_k.$$

so $\mathbf{E}_\rho |h(\rho) - h(X_1)| = o(1)$ as $k \rightarrow \infty$, but since it does not depend on k it must be zero. So $h(\rho) = h(X_1)$ almost surely. By stationarity, $h(X_n) = h(X_{n+1})$ almost surely, since h , shifted to X_n , still has a subsequence of times n_k where it is sublinear. Hence $h(\rho) = h(X_n)$ for every n , almost surely. By irreducibility, it is constant. \square

Proof of Theorem 5. Let $d \geq 2$, the constant functions on ω are obviously harmonic. The projections of $x + \chi(x)$ where χ is the corrector (see Theorem 15) on each coordinate provide us with d linearly independent functions. These functions have linear growth. Therefore, the space of linear growth harmonic functions is at least $(d+1)$ -dimensional.

Now, let h be a harmonic function on ω with (at most) linear growth and with $h(0) = 0$. Proposition 19 allows us to extract a subsequence of (h_{n_k}) converging uniformly on any compact subset of \mathbb{R}^d to a continuous function \tilde{h} . Assume for a moment that \tilde{h} is proved to be linear. Then, $f(x) := h(x) - \tilde{h}(x + \chi(x))$ is a linear harmonic function on ω with the following additional property: for every ε there is a k_0 such that for all $k > k_0$,

$$|f(x)| \leq \varepsilon n_k \quad \forall x \text{ with } |x| < \frac{1}{\varepsilon} n_k.$$

i.e. it is sublinear on a sequence of (space) scales. A simple calculation with the Gaussian upper bounds (22) shows that it is also sublinear on a sequence of time scales i.e.

$$\mathbf{E}[f(X_{n_k^2})^2] \leq \varepsilon n_k^2 \quad \forall k > k'_0.$$

By Corollary 21, it must be constant. Since $f(0) = 0$ that constant is zero and $h = \tilde{h} \circ \chi$. Therefore, any harmonic function with growth at most linear and equal to 0 at 0 belongs to a vector space of dimension d and the result follows. In conclusion, it is sufficient to prove that \tilde{h} is linear, or equivalently that it is harmonic, since it is well known that harmonic functions with at most linear growth on \mathbb{R}^d are the affine maps (take the partial derivative along one direction, it is a bounded harmonic map on \mathbb{R}^d , and thus a constant map).

Let now B_t be Brownian motion. Our first goal is to examine Brownian motion starting from 0, namely, we wish to show

$$\mathbf{E}_0[\tilde{h}(B_t)] = \tilde{h}(0) \quad \forall t > 0. \quad (29)$$

To see (29) note that h is harmonic and hence $\mathbf{E}_\rho[h(X_t)] = h(\rho)$ or equivalently

$$\mathbf{E}_0[h_{n_k}(X_{n_k^2 t}/n_k)] = h_{n_k}(0).$$

The central limit theorem (Theorem 15) allows to control $h(X_t)$ in a ball of radius $\approx \sqrt{t}$. Namely, because $(X_{n_k^2 t}/n)$ converges weakly to B_t ,

$$\left| \mathbf{E}_0\left[\tilde{h}(X_{n_k^2 t}/n_k) \cdot \mathbf{1}_{|X_{n_k^2 t}/n_k| < K}\right] - \mathbf{E}_0\left[\tilde{h}(B_t) \cdot \mathbf{1}_{|B_t| < K}\right] \right| \leq \varepsilon(K) + o(1)$$

where $\varepsilon(K)$ goes to 0 as $K \rightarrow \infty$ and the $o(1)$ is as $k \rightarrow \infty$. The Gaussian bounds (22) and the linear bound on h_{n_k} and \tilde{h} allow to control $h(X_t)$ outside that ball

$$\left| \mathbf{E}_0\left[h_{n_k}(X_{n_k^2 t}/n_k) \cdot \mathbf{1}_{|X_{n_k^2 t}/n_k| \geq K}\right] \right| \leq \varepsilon(K) \quad \forall n \text{ sufficiently large}$$

and similarly for $\tilde{h}(B_t)$. This shows (29).

We now extend (29) from 0 to all points u using Lemma 18. Let us recall that weak convergence is metrizable. For example, it is equivalent to convergence in the Lévy-Prokhorov distance metric [33, Section 11.3, especially Theorem 11.3.3]. We will not need any property of the Lévy-Prokhorov metric except that it is equivalent to weak convergence.

Now fix t , and fix also some ε and some n_0 . Consider a vertex x in the cluster to be good if the Gaussian estimates (22) hold for all $n > n_0$ and if the Lévy-Prokhorov distance between $X_{n^2 t}/n$ (started from x) and B_t (started from x/n) is smaller than ε . If n_0 is sufficiently large (depending on t and ε), the probability of x to be good will be $> 1 - \varepsilon$. For the Gaussian estimates this follows directly from (22) while for the Lévy-Prokhorov distance this follows from the equivalence of Lévy-Prokhorov convergence and weak convergence. Fix therefore n_0 to satisfy this property.

Now use Lemma 18 with the function f being $f(x, y) = \mathbf{1}_{\{x \text{ is bad}\}}$ (the y variable is simply ignored) and with some arbitrary λ . We get that for sufficiently large n , the number of bad x in $\mathbf{B}_\rho(\lambda n)$ is bounded by $C(\lambda)n^d \mathbb{P}(0 \text{ is bad}) \leq C(\lambda)\varepsilon n^d$. Define

$$B_n := \{u \in \mathbb{R}^d : |u| \leq \lambda, un \text{ is bad}\}$$

(where as usual we in fact take the point of the infinite cluster closest to un and check whether it is bad). Since the measure of B_n is $\leq C\varepsilon$, we see that, except for a set of measure $\leq C\varepsilon$ every $u \in \mathbb{R}^d$ with $|u| \leq \lambda$ satisfies that un is good for infinitely many n . But ε (both for the error and for the measure of the bad set) was arbitrary. Taking $\varepsilon \rightarrow 0$ and then $\lambda \rightarrow \infty$ we see that for almost every $u \in \mathbb{R}^d$ there is a sequence $n'_k = n'_k(u)$ (a subsequence of n_k) such that

1. The Gaussian estimates hold for $X_{n'_k}$ started from un'_k
2. The Lévy-Prokhorov distance between $X_{(n'_k)^2 t}/n'_k$ started from un'_k and Brownian motion started from u goes to zero.

Using again the equivalence of Lévy-Prokhorov convergence and weak convergence we get that random walk started from un'_k converges to Brownian motion started from u . We can now repeat the argument that led to (29) literally and get

$$\mathbf{E}_u[\tilde{h}(B_t)] = \tilde{h}(u)$$

for almost every u . Since \tilde{h} is continuous, this in fact holds everywhere. Since t was arbitrary, $\tilde{h}(B_t)$ is a continuous martingale, from any starting point.

The theorem is now proved. Using the strong Markov property we get that $\tilde{h}(u)$ is equal to its average over a sphere of arbitrary radius around u , in other words, we have established the mean-value property hence \tilde{h} is harmonic. \square

A natural extension of the supercritical bond percolation setting is to look at random environments on \mathbb{Z}^d , such as random conductance with uniformly elliptic conditions. See [67] for the existence of the corrector in this case. Similar results can probably be obtained in this setting.

5 Heat kernel derivative estimates

Our purpose in this section is to prove Theorem 6 which gives an upper bound for the (discrete) derivative of the heat kernel, $\mathbf{p}_n(x, y) - \mathbf{p}_{n-1}(x', y)$, for $x \sim x'$, where $\mathbf{p}_n(x, y) := \mathbf{P}_x(X_n = y)$.

We start with a lemma true on any graph. It relates the infinity norm of the gradient of the heat kernel to the infinity norm of the heat kernel and the entropy.

Lemma 22. *Let G be a graph of maximal degree d , for any $x, x', y \in G$ then with $x \sim x'$,*

$$\begin{aligned} (\mathbf{p}_{2n}(x, y) - \mathbf{p}_{2n-1}(x', y))^2 &\leq \\ &\leq 4d(d+1) \cdot \Delta_n(x, x')^2 \cdot \max_{\substack{a, b \in \mathbf{B}_x(2n): \\ \mathbf{d}(a, b) \geq \mathbf{d}(x, y)/2}} \mathbf{p}_n(a, b) \cdot \max_{a, b \in \mathbf{B}_x(2n)} \mathbf{p}_n(a, b) \end{aligned} \quad (30)$$

where Δ_n is defined in (6).

Proof. We know by Markov's property that

$$\mathbf{p}_{2n}(x, y) - \mathbf{p}_{2n-1}(x', y) = \sum_{a \in G} (\mathbf{p}_n(x, a) - \mathbf{p}_{n-1}(x', a)) \mathbf{p}_n(a, y).$$

Let us split the sum on $a \in G$ into two sums $I + II$, where I is the sum over $a \in \mathbf{B}_x(\mathbf{d}(x, y)/2)$, and II on the remaining a . Using Cauchy-Schwarz we can write

$$I^2 \leq \left(\sum_{a \in \mathbf{B}_x(\mathbf{d}(x, y)/2)} (\mathbf{p}_n(x, a) - \mathbf{p}_{n-1}(x', a))^2 \right) \left(\sum_{a \in \mathbf{B}_x(\mathbf{d}(x, y)/2)} \mathbf{p}_n(a, y)^2 \right)$$

For the first term, bound the denominator in the definition of Δ by its maximum and get

$$\sum_{a \in \mathbf{B}_x(\mathbf{d}(x, y)/2)} (\mathbf{p}_n(x, a) - \mathbf{p}_{n-1}(x', a))^2 \leq \Delta_n(x, x')^2 \cdot \max_{a \in \mathbf{B}_x(\mathbf{d}(x, y)/2)} \{\mathbf{p}_n(x, a) + \mathbf{p}_{n-1}(x', a)\}$$

For the second term write

$$\begin{aligned}
\sum_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y)^2 &\leq \left(\max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y) \right) \cdot \left(\sum_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y) \right) \\
&\leq \left(\max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y) \right) \cdot \left(\sum_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} d \cdot \mathbf{p}_n(y, a) \right) \\
&\leq d \cdot \left(\max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y) \right).
\end{aligned}$$

Together we get

$$I^2 \leq d \cdot \Delta_n(x, x')^2 \cdot \max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \{ \mathbf{p}_n(x, a) + \mathbf{p}_{n-1}(x', a) \} \cdot \max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y). \quad (31)$$

Now, the second maximum in the right-hand side of (31) is a maximum on a smaller set than the first maximum in (30) (note that points in $\mathbf{B}_x(\mathbf{d}(x,y)/2)$ are at distance larger than $\mathbf{d}(x,y)/2$ from y). Similarly, the first maximum is smaller than $(1+d)$ times the second maximum of (30). Therefore, the product of maxima is smaller than

$$(d+1) \cdot \max_{\substack{a, b \in \mathbf{B}_x(2n): \\ \mathbf{d}(a,b) \geq \mathbf{d}(x,y)/2}} \mathbf{p}_n(a, b) \cdot \max_{a, b \in \mathbf{B}_x(2n)} \mathbf{p}_n(a, b).$$

The estimate for II is the similar

$$II^2 \leq d \cdot \Delta_n(x, x')^2 \cdot \max_{a \notin \mathbf{B}_x(\mathbf{d}(x,y)/2)} \{ \mathbf{p}_n(x, a) + \mathbf{p}_{n-1}(x', a) \} \cdot \max_{a \notin \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y).$$

It is easy to obtain the same bound again, except the estimates are reversed (*i.e.* what was bounded by the first term before is now bounded by the second term). We sum up:

$$\begin{aligned}
(\mathbf{p}_{2n}(x, y) - \mathbf{p}_{2n-1}(x', y))^2 &= (I + II)^2 \leq 2(I^2 + II^2) \\
&\leq 4d(d+1) \cdot \Delta_n(x, x')^2 \cdot \max_{\substack{a, b \in \mathbf{B}_x(2n): \\ \mathbf{d}(a,b) \geq \mathbf{d}(x,y)/2}} \mathbf{p}_n(a, b) \cdot \max_{a, b \in \mathbf{B}_x(2n)} \mathbf{p}_n(a, b). \quad \square
\end{aligned}$$

Proof of Theorem 6. As before, percolation can be seen as a stationary random graph, and it is sufficient to prove

$$\mathbb{E} \left((\mathbf{p}_{2n}(\rho, x) - \mathbf{p}_{2n-1}(\tilde{X}_1, x))^2 \cdot \mathbf{1}_{\{x \in \omega\}} \right) \leq \frac{C'_3}{n^{d+1}} \exp(-C'_4 |x|^2/n)$$

where C'_3 and C'_4 depend only on d and the percolation probability p . Also note that one can restrict ourselves to $|x| \leq n^{1/2+\varepsilon/3}$, since in the regime $|x| \geq n^{1/2+\varepsilon/3}$, the heat kernel decreases fast enough so that one can tune the constant C'_4 in order to obtain the result for free.

Again we use the variables $n_y(\omega)$ from (1,2). Take ε to be given by the stretched exponential bound for $n_y(\omega)$, (2). Fix $|x| \leq n^{1/2+\varepsilon/3}$. Let $N(\omega) = \max\{n_y(\omega) : y \in \mathbf{B}_\rho(n)\}$.

So the Gaussian estimates (1) imply that for a.e. environment ω such that $x \in \omega$, whenever $n \geq N(\omega)$, we have

$$\max_{\substack{a, b \in \mathbf{B}_0(n): \\ \mathbf{d}(a, b) > \mathbf{d}(\rho, x)/2}} \mathbf{p}_n(a, b) \leq \frac{C_3}{n^{d/2}} \exp[-C_4|x|^2/n] \quad \text{and} \quad \max_{a, b \in \mathbf{B}_\rho(n)} \mathbf{p}_n(a, b) \leq \frac{C_3}{n^{d/2}}. \quad (32)$$

Averaging (30) on the environments satisfying $N(\omega) \leq n$ (for which we have (32)), we find

$$\begin{aligned} & \mathbb{E}[(\mathbf{p}_{2n}(\rho, x) - \mathbf{p}_{2n-1}(\tilde{X}_1, x))^2 \cdot \mathbf{1}_{\{x \in \omega\}} \mathbf{1}_{\{N(\omega) \leq n\}}] \\ & \leq 4d(d+1) \cdot \mathbb{E}[\Delta_n(\rho, \tilde{X}_1)^2] \cdot \frac{C_3^2}{n^d} \exp[-C_4|x|^2/n]. \end{aligned}$$

We now apply Theorem 7 to bound $\mathbb{E}[\Delta_n^2]$ by $2(H_n - H_{n-1})$. Recall also that Lemma 20 from the previous section says that for critical percolation $H_n - H_{n-1} \leq C/n$ for all n . Together these give

$$\mathbb{E}[(\mathbf{p}_{2n}(\rho, x) - \mathbf{p}_{2n-1}(\tilde{X}_1, x))^2 \cdot \mathbf{1}_{\{x \in \omega\}} \mathbf{1}_{\{N(\omega) \leq n\}}] \leq 4d(d+1) \frac{C \cdot C_3^2}{n^{d+1}} \exp[-C_4|x|^2/n].$$

We do not need to control the behavior of the gradient on $\{N(\omega) > n\}$ since this event has probability at most $Cn^d e^{-n^\varepsilon}$. Hence in the regime $|x| \leq n^{1/2+\varepsilon/3}$ we find

$$\begin{aligned} & \mathbb{E}[(\mathbf{p}_{2n}(\rho, x) - \mathbf{p}_{2n-1}(\tilde{X}_1, x))^2 \cdot \mathbf{1}_{\{x \in \omega\}} \mathbf{1}_{\{N(\omega) > n\}}] \\ & \leq \mathbb{P}_p(N(\omega) > n) \leq 4d(d+1) \cdot \frac{C_3^2}{n^{d+1}} \exp[-C_4|x|^2/n]. \end{aligned}$$

Putting all the pieces together, we obtain the result. \square

The proof involved only Gaussian estimates at mesoscopic scale and the entropy argument. It extends to other contexts such as random conductances satisfying the uniform elliptic condition (see Example 2.1). One may then get, using convolution, annealed second space-derivative and first time-derivative estimates for the heat kernel using the first space-derivative estimates. We refer to Section 5 of [29] for more details.

6 Open questions

This study must be understood as an introduction and some initial steps in the subject. There are many natural questions on harmonic functions which remain open. We present few of them in this section.

Minimal growth harmonic functions. The question of minimal growth harmonic functions was implicitly studied in the literature: the failure of the Liouville property corresponds to a special case of minimal growth. When the Liouville property is true, it becomes interesting to determine the minimal growth. Even the deterministic case (*i.e.* transitive or Cayley graphs) has interesting phenomenology and we plan to analyze

some examples in a future paper. Note that groups always admit linear growth harmonic functions [51, 66, 68]. This is no longer the case for stationary random graphs. When the random walk is subdiffusive (note that the random walk on Cayley graphs is at least diffusive, a result due to Erschler, see Lee and Peres [55]), Theorem 3' (page 16) implies a phenomenon which is specific to random environments:

Proposition 23. *Let (G, ν, ρ) be a stationary random graph with polynomial growth such that the random walk is (strictly) subdiffusive. Then, almost surely there do not exist linear growth harmonic functions.*

Therefore graphical fractals, UIPQ, critical Galton-Watson trees conditioned to survive and the incipient infinite cluster (IIC) do not admit linear growth harmonic functions. We mention that it was already proved [9] that the Uniform Infinite Planar Triangulation is almost surely Liouville. There are no non-constant harmonic functions on the critical Galton-Watson tree or on the IIC as both have infinitely many cut vertices. Indeed, the Galton-Watson tree is well-known to be one-ended and hence, as a tree, must have infinitely many cut vertices. The existence of cut points for the IIC is essentially known, but we did not find a reference and including a full proof would take us too far off-topic.

Question 1. Do there exist non-constant harmonic functions with polynomial growth on the UIPQ?

If such functions exist, we may ask the following question:

Question 2. What is the minimal growth of a non-constant harmonic function on the UIPQ?

Space of harmonic functions with polynomial growth. Cayley graphs with polynomial growth automatically satisfy the volume doubling property and the Poincaré inequality, thus implying that spaces of harmonic functions with prescribed polynomial growth are finite dimensional. The possibility of such behavior in the case of stationary random graphs of polynomial volume growth is a legitimate question.

Question 3. Let (G, ν, ρ) be a stationary random graph with polynomial growth. Is the space of harmonic functions with prescribed polynomial growth finite dimensional?

The difficulty comes from the fact that we do not necessarily have Poincaré inequality at our disposal (in the case of the UIPQ for instance). Therefore, we cannot use the technology developed in Section 3 in the general context. We mention that there exists another strategy to prove finite dimensionality, proposed in [27], relying on the following weaker statement: for every harmonic function h on G and $x \in G$:

$$h^2(x) \leq \frac{C}{\nu(\mathbf{B}_x(n))} \sum_{y \in \mathbf{B}(x, Cn)} h^2(y) \nu(y).$$

where C is a constant independent of x and n . This inequality appears in standard proofs of elliptic Harnack inequalities and holds for a larger class than those satisfying the doubling volume property and Poincaré inequality. Still one cannot expect it to hold in graphs with small “bottlenecks” like the UIPQ.

Question 4. Is the space of harmonic functions with some prescribed polynomial growth on the UIPQ finite dimensional?

Dimension of spaces of harmonic functions. The computation of the dimension of spaces of harmonic functions does not restrict to the case of linear growth harmonic functions. For a graph G and $k > 0$, let $d_k[G]$ be the dimension of the space of harmonic functions with growth bounded by a polynomial of degree k .

The similarity between \mathbb{Z}^d and the infinite cluster of percolation might extend to the dimension of the space of harmonic functions with arbitrary polynomial growth. More precisely, we ask the following question:

Question 5. Do the families $(d_k[\omega])_{k \geq 0}$ and $(d_k[\mathbb{Z}^d])_{k \geq 0}$ have equal dimension almost surely?

In particular, an interesting intermediate step toward this question would be to show that there is no harmonic function with non-integer growth.

It is natural to ask if an invariance principle for the random walk in the random environment ω implies that the sequence $(d_k[\omega])$ coincides with $(d_k[\mathbb{Z}^d])$. On \mathbb{Z}^d , diffusivity and the invariance principle are robust under rough isometry. Therefore, one can ask if $(d_k[G])_{k \geq 0}$ is invariant under rough isometry for these kind of graphs. This is not true in general. For instance, Liouville property is not invariant under rough isometry (see [57] for the first example or [12] for a simpler one).

More generally, one can ask whether a small perturbation of a Cayley graph modifies drastically the harmonic functions on it. For instance, consider percolation on a Cayley graph G such that $p_u(G)$ (the infimum of the values for which there exists a *unique* infinite cluster) is strictly smaller than 1. Fix $p > p_u(G)$ and set $\omega(G)$ to be the unique infinite cluster of the percolation with parameter p .

Question 6. Are the dimensions of spaces of harmonic functions with a given growth equal for G and $\omega(G)$?

Note that the question, in the case of bounded harmonic functions, was addressed in [11].

In the context of Cayley graphs, the space of harmonic functions with a certain growth rate is crucial in the study of the underlying group. Indeed, the latter acts on harmonic functions naturally. In the random setting, we do not have this interpretation. Nevertheless, an interesting question is to understand what information on the random graph is encoded in the sequence $(d_k[G])_{k \geq 0}$. In particular, the following question would be a first step in this direction:

Question 7. Consider a random subgraph G of \mathbb{Z}^d . What are the requirements to ensure that $(d_k[G])_{k \geq 0}$ equals $(d_k[\mathbb{Z}^d])_{k \geq 0}$?

Acknowledgements We wish to thank Ofer Zeitouni for pointing out that the entropy argument works also in non-reversible setting, and Russell Lyons for directing us to [34]. This paper was written during the visit of the second author to the Weizmann Institute in Israel. The first author is the incumbent of the Renee and Jay Weiss Professorial Chair.

The second author was supported by the EU Marie-Curie RTN CODY, the ERC AG CONFRA, as well as by the Swiss FNS and the Weizmann institute. The third author was supported by the Israel Science Foundation.

References

- [1] Aldous D. and Lyons R., *Processes on unimodular random networks*. Electron. J. Probab. **12**(54) (2007), 1454–1508. Available at: [washington.edu](http://www.washington.edu)
- [2] Avez A., *Harmonic functions on groups*. Differential geometry and relativity, 27–32. Mathematical Phys. and Appl. Math., Vol. 3, Reidel, Dordrecht, 1976.
- [3] Barlow M. T., *Which values of the volume growth and escape time exponent are possible for a graph?* Rev. Mat. Iberoamericana **20**(1) (2004), 1–31. [projecteuclid](http://projecteuclid.org)
- [4] Barlow M. T., *Random walks on supercritical percolation clusters*. Ann. Probab. **32**(4) (2004), 3024–3084. Available at: [projecteuclid](http://projecteuclid.org)
- [5] Barlow M. T., *Diffusions on Fractals*. St Flour Lecture Notes in Lect. Notes Math. **1690**.
- [6] Barlow M. T. and Bass R. F., *Brownian motion and harmonic analysis on Sierpinski carpets*. Canad. J. Math. **51**(4) (1999), 673–744. Available at: cms.math.ca
- [7] Barlow M. T. and Hambly B. M., *Parabolic Harnack inequality and local limit theorem for percolation clusters*. Electron. J. Probab. **14**(1) (2009), 1–27. Available at: [washington.edu](http://www.washington.edu)
- [8] Barlow M. T. and Perkins, E. A. *Symmetric Markov chains in \mathbb{Z}^d : how fast can they move?* Probab. Theory Related Fields **82**(1) (1989), 95–108. [springerlink.com](http://www.springerlink.com)
- [9] Benjamini I. and Curien N., *Ergodic theory on stationary random graphs*. Preprint (2010). Available at: [arXiv:1011.2526](http://arxiv.org/abs/1011.2526)
- [10] Benjamini I. and Curien N., *The random walk on the uniform infinite planar quadrangulation is subdiffusive*. Preprint (2011).
- [11] Benjamini I., Lyons R. and Schramm O., *Percolation perturbations in potential theory and random walks*. Random walks and discrete potential theory (Cortona, 1997), 56–84, Sympos. Math., XXXIX, Cambridge Univ. Press, Cambridge, 1999. Available at: [arXiv:9804010](http://arxiv.org/abs/9804010)
- [12] Benjamini I. and Schramm O., *Harmonic functions on planar and almost planar graphs and manifolds, via circle packings*. Invent. Math., **126**(3) (1996), 565–587. Available at: [springerlink.com](http://www.springerlink.com)
- [13] Benjamini I. and Schramm O., *Recurrence of distributional limits of finite planar graphs*. Electron. J. Probab. **6** (2001), 13 pp. Available at: [washington.edu](http://www.washington.edu)

- [14] Berger N. and Biskup M., *Quenched invariance principle for random walk on percolation clusters*. Probab. Theory Related Fields **137**(1–2) (2007), 83–120. Available at: springerlink.com, [arXiv:0503576](https://arxiv.org/abs/0503576)
- [15] Berger N. and Deuschel J.-D., *A quenched invariance principle for non-elliptic random walk in i.i.d. balanced random environment*. Preprint (2011). [arXiv:1108.3995](https://arxiv.org/abs/1108.3995)
- [16] Biskup M. and Prescott T., *Functional CLT for random walk among bounded random conductances*. Electron. J. Probab. **12** (2007), 1323–1348. Available: [washington.edu](http://www.washington.edu)
- [17] Bolthausen E., Sznitman A.-S. and Zeitouni O., *Cut points and diffusive random walks in random environments*. Ann. Inst. H. Poincaré Probab. Statist. **39**(3) (2003), 527–555. Available at: [sciencedirect.com](http://www.sciencedirect.com)
- [18] Caputo P., Faggionato A. and Prescott T., *Invariance principle for Mott variable range hopping and other walks on point processes*. Preprint (2009). [arXiv:0912.4591](https://arxiv.org/abs/0912.4591)
- [19] Carne T., *A transmutation formula for Markov chains*. Bull. Sci. Math. **109** (1985), 399–405.
- [20] Choquet G. and Deny J., *Sur l'équation de convolution $\mu = \mu * \sigma$* . [French: On the convolution equation $\mu = \mu * \sigma$.] C. R. Acad. Sci. Paris **250** (1960), 799–801.
- [21] Colding T. and Minicozzi W., *Harmonic functions on manifolds*. Ann. Math. **146**(3) (1997), 725–747. Available at: [jstor](http://www.jstor.org)
- [22] Colding T. and Minicozzi W., *Estimates for parametric elliptic integrands*. Internat. Math. Res. Notices **6** (2002), 291–297. Available at: [oxfordjournals](http://www.oxfordjournals.org)
- [23] Conlon J. G. and Naddaf A., *Green's functions for elliptic and parabolic equations with random coefficients*. New York J. Math. **6** (2000), 153–225. [albany.edu](http://www.albany.edu)
- [24] Csiszár I., *Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten*. [German: An information-theoretic inequality and its application to proving ergodicity of Markov chains] Magyar Tud. Akad. Mat. Kutató Int. Közl. **8** (1963), 85–108.
- [25] Csiszár I., *A note on Jensen's inequality*. Stud. Sc. Math. Hung. **1** (1966), 227–230.
- [26] De Giorgi E., *Sulla differenziabilità e l'analiticità del le estremali degli integrali multipli regolari*. Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. **3** (1957), 25–43.
- [27] Delmotte T., *Harnack inequalities on graphs*. Séminaire de théorie spectrale et géométrie. Grenoble. **16** (1998), 217–228.
- [28] Delmotte T., *Parabolic Harnack inequality and estimates of Markov chains on graphs*. Rev. Mat. Iberoamericana **15**(1) (1999), 181–232. Available at: [rsme.es](http://www.rsme.es)

- [29] Delmotte T. and Deuschel, J.-D., *On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model*. Probab. Theory and Related Fields **133(3)** (2005), 358–390. springerlink.com
- [30] Derriennic Y., *Quelques applications du théorème ergodique sous-additif* [French: Some applications of the sub-additive ergodic theorem]. Conference on Random Walks (Kleebach, 1979), Astérisque **74** (1980), 183–201.
- [31] Deuschel J.-D. and Kösters H., *The quenched invariance principle for random walks in random environments admitting a bounded cycle representation*. Ann. Inst. H. Poincaré Probab. Statist. **44(3)** (2008), 574–591. [projecteuclid](http://projecteuclid.org), [arXiv:0806.2525](https://arxiv.org/abs/0806.2525)
- [32] Disertori M., Spencer T. and Zirnbauer M. R., *Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model*. Commun. Math. Phys. **300(2)** (2010), 435–486. Available at: springerlink.com, [arXiv:0901.1652](https://arxiv.org/abs/0901.1652)
- [33] Dudley R. M., *Real analysis and probability*. Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics, **74**. Cambridge University Press, Cambridge, 2002.
- [34] Erschler A. and Karlsson A., *Homomorphisms to \mathbb{R} constructed from random walks*. Ann. Inst. Fourier (Grenoble) **60(6)** (2010), 2095–2113. Available: aif.cedram.org, math.kth.se/~akar1
- [35] Furstenberg H., *A Poisson formula for semi-simple Lie groups*. Ann. Math. **77(2)** (1963), 335–386. Available at: jstor.org
- [36] Furstenberg H., *Random walks and discrete subgroups of Lie groups*. In: Advances in Probability and Related Topics, **1**, 1–63. Dekker, New York, 1971.
- [37] Ferrari P. A., Grisi M. R. and Groisman P., *Harmonic deformation of Delaunay triangulations*. Preprint (2010). Available at: [arXiv:1012.1677](https://arxiv.org/abs/1012.1677)
- [38] Gloria A. and Otto F., *An optimal variance estimate in stochastic homogenization of discrete elliptic equations*. Ann. Probab. **39(3)** (2011), 779–856. [projecteuclid](http://projecteuclid.org), [arXiv:1104.1291](https://arxiv.org/abs/1104.1291)
- [39] Grimmett G., *Percolation*. Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 321. Springer-Verlag, Berlin, 1999.
- [40] Guo X. and Zeitouni O., *Quenched invariance principle for random walks in balanced random environment*. Preprint (2010). Available at: [arXiv:1003.3494](https://arxiv.org/abs/1003.3494)
- [41] Hebisch W. and Saloff-Coste L., *Gaussian estimates for Markov chains and random walks on groups*. Ann. Probab. **21(2)** (1993), 673–709. Available: projecteuclid.org
- [42] Heydenreich M., van der Hofstad R. and Hulshof T., *High-dimensional incipient infinite clusters revisited*. Preprint (2011). Available at: [arXiv:1108.4325](https://arxiv.org/abs/1108.4325)

- [43] van der Hofstad R. and Járai A. A., *The incipient infinite cluster for high-dimensional unoriented percolation*. J. Statist. Phys. **114**(3-4) (2004), 625–663. Available from: springerlink.com, tue.nl/~rhofstad
- [44] Járai A. A., *Incipient infinite percolation clusters in 2D*. Ann. Probab. **31**(1) (2003), 444–485. Available at: projecteuclid.org
- [45] David Jerison, *The Poincaré inequality for vector fields satisfying Hörmander’s condition*. Duke Math. J. **53**(2) (1986), 503–523. Available at: [projecteuclid](http://projecteuclid.org)
- [46] Kaimanovich V. and Vershik A., *Random walks on discrete groups: boundary and entropy*. Ann. Probab. **11**(3) (1983), 457–490. Available at: [projecteuclid](http://projecteuclid.org)
- [47] Kaimanovich V. and Sobieczky F., *Stochastic homogenization of horospheric tree products*. In: Probabilistic approach to geometry, **57** of Adv. Stud. Pure Math., 199–229. Math. Soc. Japan, Tokyo, 2010.
- [48] Kesten H., *The incipient infinite cluster in two-dimensional percolation*. Probab. Theory Related Fields **73**(3) (1986), 369–394. Available at: springerlink.com
- [49] Kesten H., *Subdiffusive behavior of random walks on a random cluster*. Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), 425–487.
- [50] Kipnis C. and Varadhan S. R. S., *A central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions*. Commun. Math. Phys. **104**(1) (1986), 1–19. Available at: springerlink.com
- [51] Kleiner B., *A new proof of Gromov’s theorem on groups of polynomial growth*. J. Amer. Math. Soc. **23**(3) (2010), 815–829. ams.org, [arXiv:0710.4593](https://arxiv.org/abs/0710.4593)
- [52] Kozma G. and Nachmias A., *The Alexander-Orbach conjecture holds in high dimensions*. Invent. Math. **178**(3) (2009), 635–654. springerlink.com, [arXiv:0806.1442](https://arxiv.org/abs/0806.1442)
- [53] Krikun M., *On one property of distances in the infinite random quadrangulation*. Preprint (2008). Available at: [arXiv:0805.1907](https://arxiv.org/abs/0805.1907)
- [54] Lawler G. F., *Weak convergence of a random walk in a random environment*. Commun. Math. Phys. **87**(1) (1982/83), 81–87. Available at: springerlink.com
- [55] Lee J. and Peres Y., *Harmonic maps on amenable groups and a diffusive lower bound for random walks*. Preprint (2009). Available at: [arXiv:0911.0274](https://arxiv.org/abs/0911.0274)
- [56] Li P., *The theory of harmonic functions and its relation to geometry*. Proc. Sympos. Pure Math, Amer. Math. Soc. **54**(1), (Ed. R. Greene and S.T. Yau), 1993.
- [57] Lyons T., *Instability of the Liouville property for quasi-isometric Riemannian manifolds and reversible Markov chains*. J. Differential Geom. **26**(1) (1987), 33–66. Available at: [projecteuclid](http://projecteuclid.org)

- [58] Lyons R., Pemantle R. and Peres Y., *Ergodic theory on Galton–Watson trees: speed of random walk and dimension of harmonic measure*, Ergodic Theory and Dynamical Systems, **15(3)** (1995), 593–619. Available at: cambridge.org
- [59] Lyons R. and Peres Y., *Probability on trees and networks*. Book draft. Available at: iu.edu/~rdlyons
- [60] Margulis G. A., *Positive harmonic functions on nilpotent groups* [Russian]. Dokl. Akad. Nauk SSSR **166** (1966), 1054–1057; English translation in: Soviet Math. Dokl. **7** (1966), 241–244.
- [61] De Masi A., Ferrari P. A., Goldstein S. and Wick W. D., *An invariance principle for reversible Markov processes. Applications to random motions in random environments*. J. Statist. Phys. **55(3-4)** (1989), 787–855. Available at: springerlink.com
- [62] Mathieu P. and Piatnitski A., *Quenched invariance principles for random walks on percolation clusters*. Proc. A of the Royal Soc. **463(2085)** (2007), 2287–2307. Available at: royalsocietypublishing.org
- [63] Moser J., *On Harnack’s Theorem for elliptic differential equations*. Commun. Pure Appl. Math. **14(3)** (1961), 577–591. Available at: wiley.com
- [64] Moser J., *A Harnack inequality for parabolic differential equations*. Commun. Pure Appl. Math. **17(1)** (1964), 101–134. Available at: wiley.com
- [65] Nash J., *Continuity of solutions of parabolic and elliptic equations*. Amer. J. Math. **80** (1958), 931–954.
- [66] Shalom Y. and Tao T., *A finitary version of Gromov’s polynomial growth theorem*. To appear in Geom. Func. Anal. (2011). Available at: arXiv:0910.4148
- [67] Sidoravicius V. and Sznitman A.-S., *Quenched invariance principles for walks on clusters of percolation or among random conductances*. Probab. Theory Related Fields **129(2)** (2004), 219–244. Available at: springerlink.com
- [68] Tao T., *A proof of Gromov’s theorem*. Blog (2010), terrytao.wordpress.com
- [69] Varopoulos N. Th., *Long range estimates for Markov chains*. Bull. Sci. Math. **109** (1985), 225–252.
- [70] Woess W., *Random walks on infinite graphs and groups—a survey on selected topics*. Bull. London Math. Soc. **26(1)** (1994), 1–60. Available from: oxfordjournals.org, tugraz.at/~woess
- [71] Yau S. T., *Harmonic functions on complete Riemannian manifolds*. Commun. Pure Appl. Math. **28(2)** (1975), 201–228. Available at: wiley.com

UNIVERSITÉ DE GENÈVE
GENÈVE, SWITZERLAND
E-MAIL: hugo.duminil@unige.ch
WEIZMANN INSTITUTE
REHOVOT, ISRAEL
E-MAIL: itai.benjamini@weizmann.ac.il ; gady.kozma@weizmann.ac.il
BEN GURION UNIVERSITY
BEER SHEVA, ISRAEL
E-MAIL: yadina@bgu.ac.il